

**A framework to introduce flexibility
in crop modelling:**

**from conceptual modelling
to software engineering and back**

Myriam ADAM

Thesis committee

Thesis supervisors

Prof. dr. ir. H. Van Keulen
Professor at the Plant Production Systems Group
Wageningen University

Prof. dr. J. Wery
Professor of Agronomy at Montpellier SupAgro
Director of “Unité Mixte de Recherche” System #1230
CIRAD-INRA-Supagro, Cedex, France

Thesis co-supervisors

Prof. dr. F.A. Ewert
Professor of Crop Science
Institute of Crop Science and Resource Conservation (INRES)
University of Bonn, Germany

Dr. ir. P.A. Leffelaar
Associate Professor at the Plant Production Systems Group
Wageningen University

Other members

Prof. dr. ir. G. Van Straten, Wageningen University
Prof. dr. H.B.J. Leemans, Wageningen University
Dr. ir. H. Hengsdijk, Plant Research International, Wageningen
Dr. H. Scholten, Wageningen University

This research was conducted under the auspices of the C.T. De Wit Graduate School of Production Ecology and Resource Conservation.

**A framework to introduce flexibility
in crop modelling:**

from conceptual modelling
to software engineering and back

Myriam ADAM

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University

by the authority of the Rector Magnificus

Prof. dr. M.J. Kropff

in the presence of the

Thesis Committee appointed by the Academic Board

to be defended in public

on Friday 22 October 2010

at 1:30 p.m. in the Aula

Myriam ADAM

A framework to introduce flexibility in crop modelling: from conceptual modelling to software engineering and back

190 pages

Thesis, Wageningen University, Wageningen, NL (2010)

With references, with summaries in English, Dutch and French

ISBN: 978-90-8585-764-8

ABSTRACT

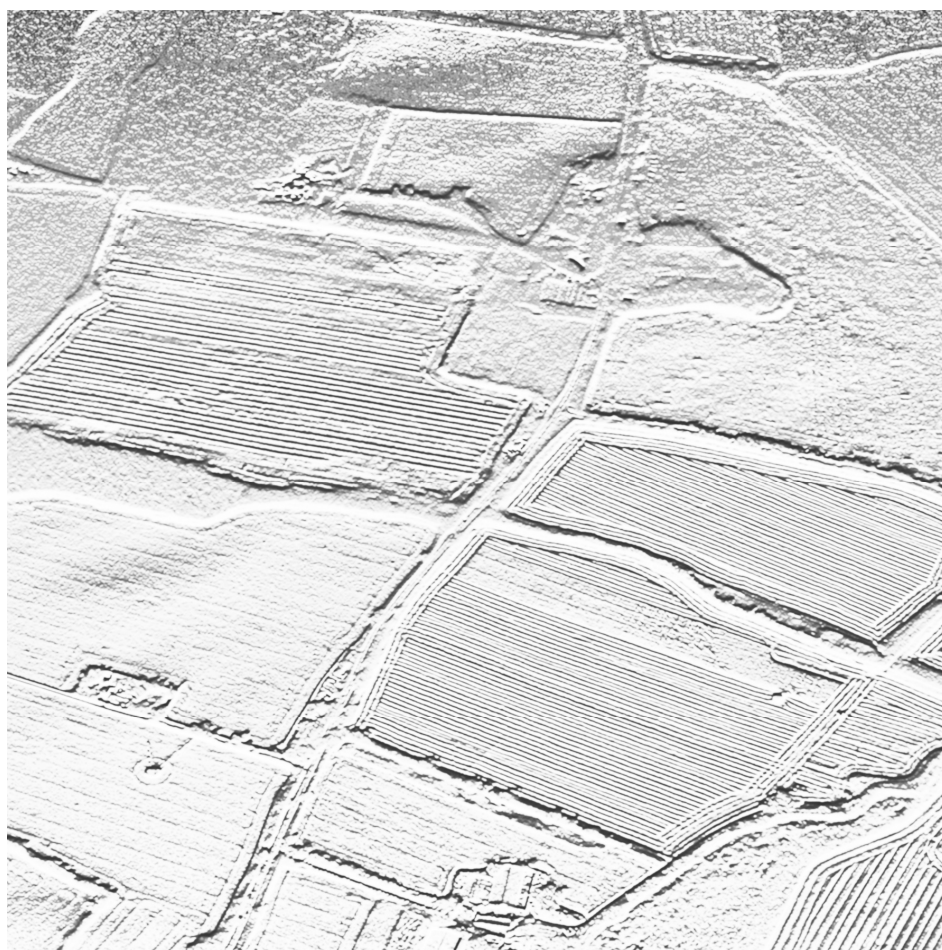
This thesis is an account of the development and use of a framework to introduce flexibility in crop modelling. The construction of such a framework is supported by two main beams: the implementation and the modelling beam. Since the beginning of the 1990s, the implementation beam has gained increasing attention in the crop modelling field, notably with the development of APSIM (Agricultural Production Systems sIMulator) in Australia, OMS (Object Modelling System) in the United States, and APES (Agricultural Production and Externalities Simulator) in Europe. The main focus of this thesis is on the modelling beam and how to combine it with the implementation beam. I first explain how flexibility is adopted in crop modelling and what is required for the implementation beam of the framework, namely libraries of modules representing the basic crop growth and development processes and of crop models (i.e. modelling solutions). Then, I define how to deal with this flexibility (i.e. modelling beam) and more specifically I describe systematic approaches to facilitate the selection of the appropriate model structure (i.e. a combination of modules) for a specific simulation objective. While developing the framework, I stress the need for better documentation of the underlying assumptions of the modules and of the criteria applied in the selection of these modules for a particular simulation objective. Such documentation should help to point out the sources of uncertainties associated with the development of crop models and to reinforce the role of the crop modeller as an intermediary between the software engineer, coding the modules, and the end users, using the model for a specific objective. Finally, I draw conclusions for the prospects of such a framework in the crop modelling field. I see its main contribution to (i) a better understanding in crop physiology through easier testing of alternatives hypotheses, and (ii) integrated studies by facilitating model reuse.

Keywords: model structure, uncertainty, modularity, software design patterns, good modelling practices, crop growth and development.

Abstract	v
Chapter 1 General Introduction	1
Chapter 2 Software engineering techniques: the implementation beam of the framework	11
Chapter 3 Systems analysis: the modelling beam of the framework	35
Chapter 4 Criteria of selection: data availability and scale	61
Chapter 5 Criteria of selection: crop type	89
Chapter 6 General Discussion	117
References	133
Appendices	147
Summary	159
Samenvatting	165
Résumé	171
Glossary	177
Acknowledgements	179
Curriculum vitae	181
List of publications	183
PE&RC education Statement form	187
Funding	190

CHAPTER 1

GENERAL INTRODUCTION



1. Crop modelling, the Wageningen crop models and their (re)use

Modelling is a way to simplify a system. A system is defined as a part of reality that contains interacting components. Crops are populations of plants of which the growth is managed by humans for any of various uses they may have for (parts of) these plants. Crop modelling is a way to simplify the part of reality (i.e. system) known as a crop.

At first, crop models were mainly developed to increase understanding of the basic processes of crop growth and development. Taking the crop models developed at Wageningen University over the years as a reference, good examples of such models are ELCROS (Elementary CROp growth Simulator, De Wit et al., 1970), BACROS (Basic CROp growth Simulator, De Wit, 1978), and SUCROS (Simple and Universal CROp growth Simulator, Van Keulen et al., 1982). Subsequently, the focus in modelling became more application-oriented and models were developed to increase understanding of the degree to which the main biotic and abiotic factors constrain crop yields. Among the Wageningen crop models, an example of such a model is WOFOST (World FOod STudies, Van Keulen and Wolf, 1986), which is still used for yield forecasting in the European Union (e.g. in the MARS project, Monitoring Agricultural ResourceS, <http://mars.jrc.it/>).

The increasing number of applications resulted in the development of a multitude of different models, and soon the need was felt to combine different models in one single framework for handling the modelling and analysis needs of different cropping systems in different environments. Two examples of such simulation frameworks are DSSAT (Decision Support System for Agrotechnology Transfer, Jones et al., 2003) and APSIM (Agricultural Production systems SIMulator, Keating et al., 2003). They provide structures to easily incorporate new models and to enable the simulation of different crops. So far, no such framework can be identified within the pedigree of crop growth simulation models of Wageningen's "School of De Wit" (Bouman et al., 1996; Van Ittersum et al., 2003).

Finally, the emphasis shifted towards addressing the need to create generic crop modelling tools to support scientific investigations and facilitate decision-making for crop managers (Hammer et al., 2002). Generic models are based on common model algorithms and structure to simulate 'all' crops, with differences among crops being reflected in the use of different crop-specific sets of parameters. Among the crop growth simulation models of the "School of De Wit", examples of such a model are GECROS (Genotype-by-Environment interaction CROp growth Simulator, Yin and Van Laar, 2005) or WOFOST (Van Keulen and Wolf, 1986).

Diversity of models leads to a diversity of reuses and abuses of models (Boote et al., 1996). Models are used to describe, simulate, extrapolate and ultimately understand the function of dynamic systems. This is not different for crop models: they can be, and are, used for various objectives (from understanding crop functioning to scenario analysis exploration) by different users (from researchers to policy makers) and at various scales (from gene to globe). Because of this plethora of models and their uses, the quest for balance in crop modelling between the objectives of the simulation and the approaches selected has become a subtle exercise. It involved navigating between very detailed modelling approaches, aiming at process understanding for a wide range of conditions rather than at yield prediction, and very empirical approaches primarily aiming at prediction and specific to a particular location (Monteith, 1996; Rastetter et al., 1992).

As a response to this quest for a better balance, emphasis has been put on the development of frameworks which support flexible model composition, the composition depending on the problem at stake. Progress in computer science and software engineering [e.g. object-oriented programming (Reynolds and Acock, 1997) and, more recently, component-oriented programming (Qureshi and Hussain, 2008)] supports the development of modular approaches to create such modelling frameworks that serve as vehicles to assemble models according to different modelling aims. As a result, in the past 10 years, modular frameworks have been promoted as a way (i) to carry out comparative studies, in which they are viewed as playing a heuristic role, and (ii) to integrate crop models in systems with broader boundaries, by building on crop models already existing capability (Hammer et al., 2002). In the Wageningen crop models, the development of FSE (Fortran Simulation Environment, Van Kraalingen, 1995) and FST (Fortran Simulation Translator, Rappoldt and Van Kraalingen, 1996) can be considered attempts for creating such frameworks. Moreover, Wageningen's recent involvement in the development of APES (Agricultural Production and Externalities Simulator, Donatelli et al., 2010) within the SEAMLESS project (Van Ittersum et al., 2008) is an example of this new emphasis. Development of such modular frameworks creates flexibility in crop modelling. But how can we deal with modularity and flexibility in crop modelling, and how can we create balanced composition?

2. Modularity and flexibility in crop modelling: implementation level

Modularity is the property of a system to be made up of relatively independent, but interacting components or parts. In crop modelling, it refers to the possibility to assemble a model from different components and/or modules representing crop

growth and development processes and soil water and nitrogen dynamics. The extent to which a system is broken down into small parts (e.g. component and modules) is called granularity. A *component* is a piece of software used for composition of a model, while a *module* is a conceptualization of a process implemented within a component (e.g. radiation use efficiency or photosynthesis to calculate biomass production).

Flexibility refers to the quality of being adaptable. In crop modelling, flexibility is given through the ability to create different modelling solutions (MS) on the basis of the specific objective of the simulation study. A *modelling solution* (i.e. crop model) is the result of the combination of different modules and components, selected in dependence of the system studied (a crop model adapted to a given crop, in a given environment and for a given question).

Granularity (i.e. level of decomposition) was also advocated by Zeigler and Marshall (1991) in the development of the Discrete Event System Specification (DEVS). Figure 1.1 illustrates the different levels of granularity introduced above in comparison with the terminology from Zeigler et al. (2000). The DEVS is based on a collection of models to be assembled to satisfy a new requirement (system of systems, SoS), and these models are an assemblage of what they called primitives (Figure 1.1).

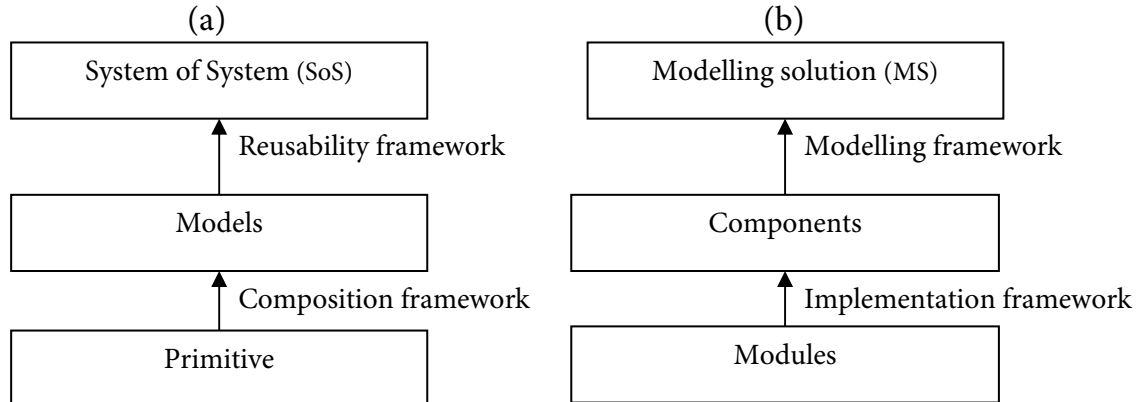


Figure 1.1. Modularity and flexibility are represented via different levels of granularity within (a) DEVS vs. (b) APES and this thesis.

Accordingly, flexibility and modularity in crop modelling result in the construction of building blocks with different levels of granularity, representative for different parts of the systems under study. The construction of these building blocks is facilitated by what Van Evert et al. (2005) called an “implementation-level” framework, the main purpose of which is to link existing models, often treated as black boxes. Subsequently, the main research questions arising from modularity and flexibility are: (i) how do we make existing models work together; and (ii) how do we design new modelling solutions for easier and balanced assembly?

3. Definition of concept for model construction: modelling level

Model conceptualization remains the heart of the matter (Donatelli et al., 2002), and there is a need to couple principles of systems analysis with the application of new software engineering techniques (i.e. use of design patterns, component-oriented design) in crop modelling. These current software engineering techniques enable an easy plug-and-play of modules (Papajorgji, 2005). However, such an easy assemblage of modules needs to be associated with guidance for model adaptation/re-assembly for the system studied (a given crop, in a given context, and for a given question). This guidance should be provided by a structured and coherent set of concepts to (i) capture similarities and differences among crops and (ii) define the basic properties of the cropping system. These concepts can be translated into selection criteria or systematic approaches to guide the user of the framework in the selection of modules or components, to assemble a new MS.

A concept can be defined as the “logic to assemble of the appropriate modules”. The development of such a concept should be based on systems analysis principles, enabling to go from the objective of the simulation to the right assembly of modules, representative of the Genotype \times Environment \times Management interactions of the cropping system under study (Figure 1.2). The formulation of concepts supports the construction of the model structure, in agreement with the users’ simulation objective. Indeed, these concepts contextualize the model for a specific application (Villa, 2007) by (i) arriving at an accurate and unambiguous problem definition, (ii) identifying the relevant factors to consider for construction of the model, and finally (iii) defining the model structure, i.e. the modelling solution.

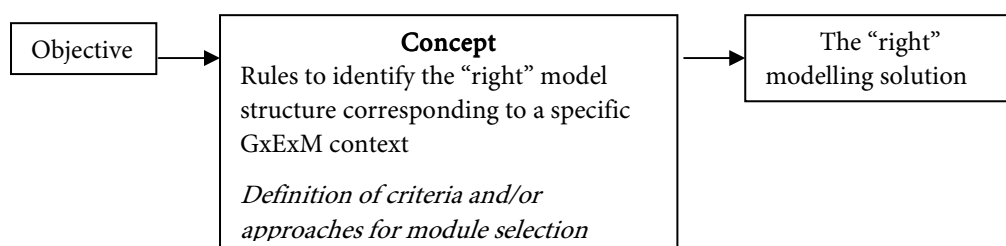


Figure 1.2. From objective(s) to the “right modelling solution”: construction of a concept to guide the user in module selection.

Formulating concepts is achieved mainly through comparison of the various existing crop modelling approaches (i.e. modules) and on the basis of a thorough understanding of crop physiology and the user’s demand. Concepts represent this understanding and should, therefore, correspond to the main criteria guiding the selection of a specific modelling solution (Figure 1.2). To this end, those concepts express rules to identify the specific combinations of modules needed for the

simulation objective. Formulation of these concepts produces an explicit analysis of the trade-offs in terms of level of details and generality among crop modelling approaches. The resulting criteria and systematic approaches can be considered as the “modelling-level” framework (Van Evert et al., 2005), aiming at guiding the user in the selection of modules and the assembly of the modelling solution. To define these criteria and to use systematic approaches helping to go from the simulation objective to the “right modelling solution” (Figure 1.2.), various type of experts are involved among which are crop physiologists, agronomists, crop modeller, and software engineers.

4. Objective

Given this context, a project was launched to develop an approach that should facilitate model (re-)use, by integrating new software engineering techniques in crop modelling (Donatelli et al., 2002) and system analysis, to enable to build models from (physiologically) meaningful components (Ewert, 2008). This project was initiated within the SEAMLESS project (Van Ittersum et al., 2008) to re-enforce collaboration between the various developers of the crop component of the Agricultural Production and Externalities Simulator (APES). The key idea of this collaboration was the development of a modular approach that would allow exchange of models (or parts of models, i.e. modules), substitution of processes, depending on the user demand and objective (which may vary from local to regional and larger scale). This thesis focuses more specifically on the definition of criteria or approaches to guide the selection of modules to assemble a modelling solution for specific applications (e.g. exploration of management options at the whole-farm scale; integrated assessment of agricultural systems at regional scale; assessment of climate change impact on crop productivity at global scale).

The general **objective** of this thesis is the development of a framework to assemble different crop models depending on the crop system to analyse and the simulation objective. This framework is operational at the two main levels explained before: the implementation level and the modelling level. After briefly describing the main principles used for the implementation level (Chapter 2), the remainder of the thesis focuses on the modelling level. At the modelling level, two aspects receive particular attention, i.e. identification and definition of physiological similarities and differences among crops and explicit description of the context of the simulation. From this general objective, more specific objectives are derived:

- Incorporate different crop growth and development modelling approaches in the framework (implementation level);

- Explicitly formulate the hypotheses underlying different simulation approaches and their validity domains (modelling level);
- Develop systematic approaches to select the relevant modules to build a modelling solution that meets a specific simulation objective (modelling level).

5. Outline

Creating a “generic” crop modelling framework for design and assessment of innovative cropping systems requires the development of a *modular and flexible* structure associated with concepts that will help modellers and end-users to select modules and to construct the crop model (i.e. modelling solution) they need for their objective. In this thesis (Figure 1.3), Chapter 2 describes how to bring the modularity into the framework, and how new software techniques can be applied in crop modelling. From the software design presented in Chapter 2, two distinct libraries can be distinguished in the framework: one including the different modelling approaches (i.e. modules describing the basic crop processes), and one including the result of the assemblage of these modules into a new modelling solution (i.e. a crop model for a given crop, in a given context, for a given question).

Chapter 3 tackles the issue of flexibility and defines a methodology (guideline) to select and combine these different approaches into a modelling solution according to specific criteria (e.g. physiological specificities of the crop, data availability, level of detail of the modules, limiting factors of production). This methodology emphasizes the importance of explicitly formulating the hypotheses underlying different simulation approaches and their validity domains. From the principles laid out in Chapters 2 and 3, it is concluded that the framework is built from two main beams: (i) the implementation beam and (ii) the modelling beam (Figure 1.3). However, there is a need to (i) further investigate on the development of concepts for the selection of modules, and more specifically (ii) refine the criteria guiding this selection. These are the subjects of Chapters 4 and 5 (Figure 1.3).

In Chapters 4 and 5, we used the principles discussed and elaborated previously, and studied two specific case studies, developing a set of concepts to select the relevant modules to build a modelling solution that meets a specific modelling objective. Although the two chapters do not explicitly apply the whole methodology presented in Chapter 3, we can see them as an example of how to operationalise the part of the framework on the selection of modules (i.e. modelling beam).

Chapter 4 deals with the scale and data availability criteria, focussing on an analysis of the effects of modelling detail in simulating crop productivity under a wide range of climatic conditions. Through comparison of various modelling approaches (modules),

we identified the importance of clarifying the assumptions underlying the quantification of a parameter value, as a parameter may represent a different meaning (and hence a different numerical value), depending on the approach selected. We also demonstrated that when addressing issues at global scale with rather scanty data, one needs to be aware of the potential consequences of the simplification of processes, as this may lead to the omission of important relationships.

Chapter 5 focuses on crop physiology and more specifically on the “crop type” criterion. It emphasizes the importance of integrating crop physiologist’s knowledge in all the modelling processes when crop type-specific modules are selected. In practice, it illustrates an approach that supports the integration of crop physiological knowledge into the framework. It emphasizes that the modeller should have a thorough understanding of the conceptual model and should communicate/interact efficiently with software engineers in developing a new modelling solution for a new crop.

Chapter 6 synthesizes the main outcomes of the previous chapters, highlighting what is needed for the development of a framework to introduce flexibility in crop modelling and what the potential future of such a framework could be.

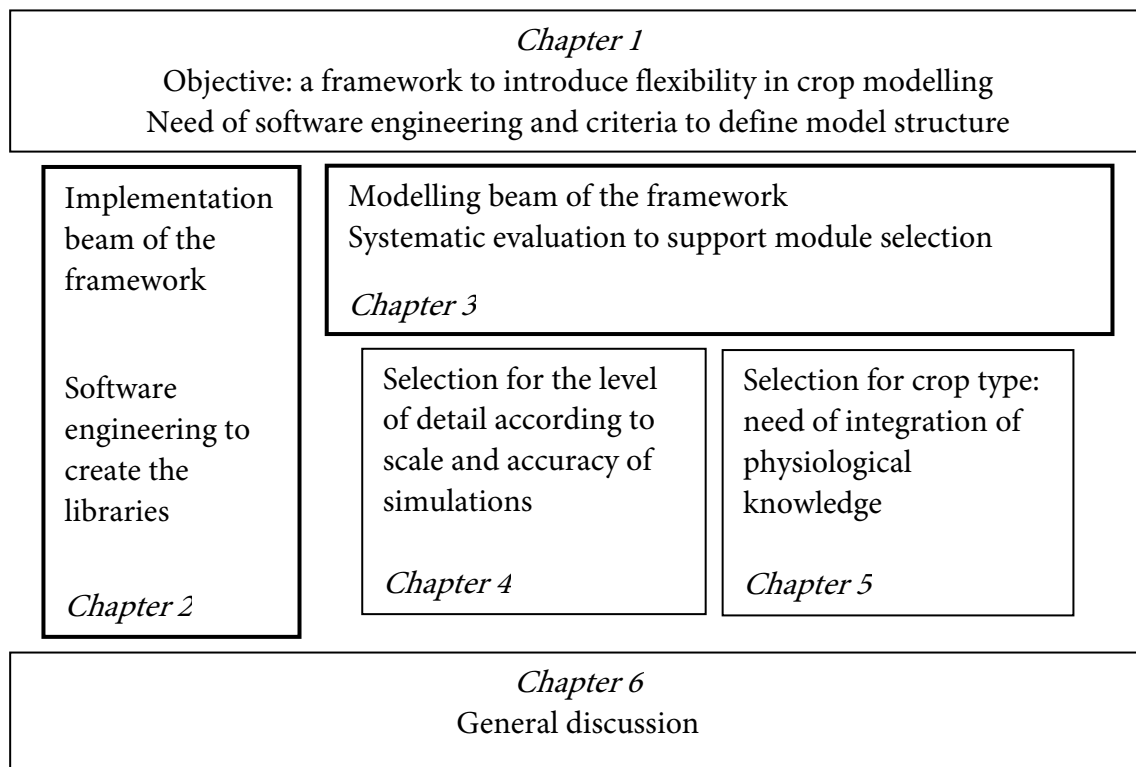


Figure 1.3. Schematic outline of this dissertation, based on two main beams: (i) the implementation beam, defined by software engineering techniques and (ii) the modelling beam, defined by an approach supporting the selection of modules. Two specific case studies refine the criteria for modules selection, considering the (i) scale and accuracy of simulations, and (ii) crop type.

CHAPTER 2

Software engineering techniques: the implementation beam of the framework*



* This chapter has been published as:

Adam, M., Ewert, F., Leffelaar, P.A., Corbeels, M., Van Keulen, H., Wery, J., 2010. CROSPAL, software that uses agronomic expert knowledge to assist modules selection for crop growth simulation. *Environmental Modelling & Software* 25, 946-955.

Abstract

Crop growth models are used for a wide range of objectives. For each objective a specific model has to be developed, because the reusability of a model is often limited by the necessity of a fundamental restructuring to adapt it to a different objective. To overcome this limitation, we developed a method to facilitate model restructuring by a novel combination of software technology with expert knowledge.

This resulted in the decision-making software application CROSPAL (CROp Simulator: Picking and Assembling Libraries). CROSPAL includes (i) a library of processes each containing different modelling approaches for each crop physiological process and (ii) a procedure based on expert knowledge of how to combine the different processes for the objective of the simulation.

A brief overview of the state of the art in crop modelling is presented, followed by an account of the developed concept to improve flexibility in crop modelling considering expert knowledge. We describe the design of the software and how expert knowledge is integrated. The use of CROSPAL is illustrated for the modelling of crop phenology. We conclude that CROSPAL is a helpful tool to improve flexibility in crop modelling considering expert knowledge but further development and evaluation is required to extend its range of application to more processes and issues crop modelling is presently addressing.

Keywords: crop models, design pattern, phenology, model flexibility, physiological processes, libraries.

1. Introduction

Crop growth modelling started in the late 1960s with De Wit (1965). Initially, crop models were developed to increase understanding of the mechanisms (i.e. physiological processes) underlying crop growth and development. Subsequently, the focus in crop modelling became more application-oriented and models were used to increase the understanding of the degree to which the main abiotic factors (weather and soil) constrain crop yields (Van Ittersum and Rabbinge, 1997). As a result, each crop growth model tended to focus on one crop and one specific set of conditions. Different modelling approaches have been developed to simulate the same process. For instance, for the simulation of biomass accumulation, various algorithms have been used. Some models comprise a very detailed description of the processes related to photosynthesis and respiration, while others use the radiation use efficiency approach, representing the detailed photosynthesis and respiration models by one parameter. Similar differences in modelling approaches and detail can be found for other crop physiological processes. Therefore, the quest for balance in crop modelling has become a subtle exercise, navigating between very detailed modelling approaches aiming at process understanding for a wide range of conditions, and empirical approaches primarily aiming at prediction, but specific to a particular set of (local) biophysical conditions (Monteith, 1996; Rastetter et al., 1992).

The crop modelling community, being aware of the specificity of the many models developed, and searching for greater generality, initiated the development of more generic models. Such generic models aim at widening the range of applicability of crop simulation models. Some examples are WOFOST (Supit et al. 1994), CROPGRO (Jones et al., 2003), STICS (Brisson et al., 2003), and EPIC/CROPSYST (Stockle et al., 2003). Reflecting on these efforts, Brisson et al. in Wallach et al. (2006: p.261) stated that “the generic nature of a model does not preclude crop specificity, but it is indicative for the efforts being made towards a common approach based on agro-physiology”. Common model algorithms and structure are adopted for all crop species and differences among species are defined in species-specific sets of parameters. Although each of these models has been widely applied, they still strongly reflect a specific modelling community (e.g. WOFOST, one of the Wageningen models; STICS, a French model; EPIC/CROPSYST, a model developed by the American crop modeller’s community). The degree to which models have been applied outside the community of the developers, can mainly be attributed to the accessibility of the specific model and the efforts to disseminate the model by the team that was involved in its original development.

Furthermore, as modelling became more application-oriented in the 1990s, attempts were made to combine different models to analyze cropping systems in different environments and with different management practices. DSSAT (Decision Support System for Agrotechnology Transfer, Jones et al., 2003) and APSIM (Agricultural Production Systems sIMulator, Keating et al., 2003) are two prominent examples of such simulation frameworks. The main objectives of these models are (i) to simulate productivity of cropping systems, taking into account weather, crop genetics, soil water, soil carbon and nitrogen dynamics, and management in single or multiple seasons and in crop rotations at any location, a common feature shared by all crop models and (ii) to provide a modular structure to facilitate incorporation of new modules - a specific objective of DSSAT and APSIM. Interaction among components takes place at the cropping system level where crop, soil and climate are identified as the key sub-systems. Thus, although these cropping system models apply principles of modularity, their structure does not allow an easy plug-in pull-out of lower level (e.g. plant process level) modules for different simulation objectives. The crop growth and development module essentially constitutes the main entity. The mechanistic detail of the physiological processes considered within the crop module cannot be easily replaced by another formulation. The crop module runs as a whole and within this module, processes are inseparable. For instance, substituting the phenology module implemented within APSIM by an alternative module (developed and required for another application) is not straightforward.

To overcome this limitation, Wang et al. (2002) proposed a generic crop model template for APSIM, based on four main components: (i) a standard crop interface, (ii) a generic crop model structure, (iii) a crop process library and (iv) a crop parameter file. The crop parameter file enables to switch between different modelling approaches, represented in the crop process library. Principles applied in this framework were already mentioned in the late 1990s when Hammer (1998) described a tree structure of a generic crop template, enabling to switch between optional processes within sub-modules. Later, this framework was applied by Van Oosterom et al. (2006) substituting the module of floral initiation with a gene network module. Such work is essential for further development of a crop modelling framework to facilitate the change of the structure of crop models.

However, despite the merit of addressing modelling choices at the level of physiological processes, no clear evidence of consistent substitution is presented. The framework facilitates substitution in terms of software engineering, but does not facilitate the user to ensure that a new method substituting another is compatible to

the modelling approaches of the other physiological processes to which it will be linked. This is particularly important for the temporal resolution of processes to avoid stiff systems.

Thus, the exchange of a modelling approach for a specific physiological process may require changes in the modules of other crop processes that may not be self-evident and may complicate the plug-in pull-out procedure. Also, the links among key physiological processes may become unbalanced (Monteith, 1996) and not suitable to the objective of the simulation. Although the framework presented by Wang et al. (2002) is a step towards the facilitation of including various modules representative of various crop growth processes, the issue of consistency in substituting one module by another within cropping system models has not been explicitly addressed yet. It remains to clarify how to identify and “evaluate the alternatives” (Hammer, 1998, p.30). Day (2001, p.217) stated “Isolating the critical processes is the first step. Describing them at an appropriate level of detail, preferably only one or two levels below that of the output of interest, is the next target.” Similarly, Acock and Acock (1991) pointed out the difference in the level of detail needed if the main objective is prediction or understanding of the mechanism representing processes in the crop system (“up to two levels of organization below the level of prediction”, p.56).

In crop modelling, the Soil-Plant-Atmosphere System Simulation (SPASS, Wang and Engel, 2000), the object-oriented crop model by Acock and Reddy (1997), the generic crop model template in APSIM (GCROP, Wang et al., 2002), the Agricultural Production and Externalities Simulator (APES, Donatelli et al., 2010) and the Common Modelling Framework of Moore et al. (2007) are illustrative examples of the few attempts to apply modularity to the simulation of the essential processes common to crop development and growth. However, explicit procedures are lacking for selecting the appropriate processes depending on the simulation objectives and defining the relationships among these processes which we specifically address in this study. Combining principles from Wang et al. (2002), object-oriented design and the use of hierarchical configuration adopted by Moore et al. (2007), we designed CROSPAL (CROp Simulator: Picking and Assembling Libraries), a crop modelling framework for the user to pick and assemble the crop simulator he/she needs. Not only should the choice of the level of detail be governed by quantitative criteria such as characteristic time of the models but also by the definition of criteria representing the objective and the targeted output of the simulation.

2. CROSPAL design: concept for a crop modelling framework

In the proposed crop modelling framework, we apply the principle of modularity to facilitate the re-use of crop models and modules for individual processes. We include modules with different descriptions of the key plant physiological processes for the user to compose the crop model (i.e. crop simulator) that meets best the simulation objective. The crop modelling framework guides the user in the selection of these descriptions to support consistency in the considered processes linked to a crop simulator. Through principles of systems analysis, comprising among others the definition of system boundaries, simplification of reality, problem definition, and identification of attributes of interest leading to the definition of criteria, the appropriate combination of modelling approaches is defined for a specific simulation objective. The crop modelling framework comprises (i) a library of plant process descriptions with different modelling approaches for each process and (ii) a procedure or workflow describing how to combine the different processes for an application, resulting in (iii) a library of crop simulators (i.e. crop models developed for specific applications).

To build the framework, we need (i) to define a set of basic crop growth and development processes according to our understanding of the crop system, (ii) to provide different modelling approaches for each process, (iii) to express expert knowledge in the form of explicit criteria to ensure consistency in the selection and linkage of modelling approaches for a specific application and (iv) to use object-oriented design to provide flexibility to the framework (Donatelli et al., 2006 a,b).

2.1. Library of crop growth and development processes

Wery (2005) defined the major crop growth and development processes as leaf area expansion, production of assimilates, partitioning of assimilates, phenology (vegetative and reproductive development), nitrogen dynamics, and transpiration. This representation of crop physiology is supported by Hay and Porter's (2006) description of the physiology of crop yield where they identified "the interlinked processes that form the basis of crop growth and yield" as phenological development, leaf canopy development and biomass production and its partitioning over the plant organs. This categorization gives information on growth and development of the crop in response to any change in the environment of the system. Each of these basic crop growth and development processes can be modelled in various ways and with different mechanistic details (Table 2.1), resulting in a variety of crop models.

Table 2.1. Proposed crop growth and development processes and selected corresponding crop modelling approaches (adapted from Wery, 2005). For comparison, the template proposed by Wang et al. (2002) is also included.

Plant growth and development processes		Modelling Approaches
Proposed structure	Related 7 main components in GCROP * (Figure 1 in Wang et al., 2002)	
Light interception: leaf area expansion, leaf canopy development	C: canopy E: senescence	<ul style="list-style-type: none"> • Development stage dependent • Biomass accumulation dependent
Dry matter production	B: biomass (B1) Partly D: root system (root biomass)	<ul style="list-style-type: none"> • Water use efficiency (WUE) • Radiation use efficiency (RUE) • Minimum (RUE, WUE) • Farquhar approach: Photosynthesis, Assimilation, Respiration
Phenology: vegetative and reproductive development	A: phenology	<ul style="list-style-type: none"> • Leaf appearance rate based on photo-thermal time • Gene network model • Simple regression
Partitioning/allocation: development of sink and assimilate partitioning	B: biomass	<ul style="list-style-type: none"> • Predetermined allocation • Source/ sink
Production level: e.g. Water stress	Partly D: root system (root depth) F-G: water and nitrogen	<ul style="list-style-type: none"> • Stomatal conductance • Transpiration, water uptake • Nitrogen uptake

* Letters A-G refer to the essential processes of a generic crop model template (Wang et al., 2002).

The library of these different modelling approaches for each individual crop growth and development processes is constructed using principles from object-oriented design to support flexibility for model development. The next step in the construction of the crop modelling framework is to define rules to help the user to select and combine the appropriate modelling approaches of these physiological processes according to the simulation objective.

2.2. Procedure to combine the modelling approaches: building a library of crop simulators

In the crop modelling framework, we aim at explicit identification of the most important characteristics of the system and how these will influence the selection of the modelling approaches to build the crop simulator. The key relationships among the different processes represented in the system are based on the knowledge of the expert and his/her understanding of the functioning of the system.

Crop growth and development is the result of the interaction of (i) the availability of resources, (ii) the ability of the crop to capture these resources and convert them into biomass and (iii) the way management of the system affects 1 and 2. Accordingly, to select the appropriate combination of modelling approaches, we apply five criteria: (i) crop physiology, (ii) limiting factors, (iii) scale (temporal and spatial), (iv) management practices, and (v) accuracy/adequacy of simulation. These criteria represent the main factors determining the simulation of crop performance. They consider Genotype x Environment x Management interactions (Yin et al., 2004), including effects of technological changes (Ewert, et al., 2005), scale and data availability.

To represent these criteria and guide the user in constructing or selecting a consistent crop simulator, we designed a graphical user interface (GUI) that helps to define the combination of the different modelling approaches. We established the following hierarchy. First, we select the modelling approach depending on crop physiology (crop type), limiting factors and scale (Figure 2.1a). Such criteria will guide the selection of the main basic crop processes to include. For example, if the user wants to simulate potential crop yield of winter wheat at field level, the choice of 'winter crop' in the GUI will direct him to pick a photo-vernal thermal modelling approach for phenology. Second, according to the choices made, another more detailed window appears to select the management and the required accuracy of the simulation (Figure 2.1b). On the basis of these selections, we retrieve a pre-defined crop model structure (crop simulator), corresponding to the user's objective. The definition of these criteria and how they influence the selection of the modelling approaches is the result of expert knowledge embedded into CROSPAL.

a. First set of criteria to decide upon

b. Second set of criteria

Different screens may appear depending on the choice made on a. We present the most complete one.

Figure 2.1. The graphical user interface (GUI) of the crop modelling framework (CROSPAL) in two main windows to illustrate the hierarchy of criteria in the choice of the model structure.

In this way the framework provides (i) pre-defined crop models to be selected from a library of crop simulators depending on the application (these crop simulators encapsulate the expert knowledge from crop modellers and software engineers), and (ii) guiding rules for the more advanced user, to develop crop models (bottom up approach, picking and assembling from the library of crop modelling approaches), extending the expert knowledge in CROSPAL. The pre-defined models can directly be run through the GUI, while the rules are the result of the process of going through the criteria guiding the user to characterize the level of modelling detail needed (Chapter

3). Once the user (a crop modeller or agronomist) has defined this level of detail, not only will he/she use the modelling approaches already included, but also the user or an associated software engineer may create the missing pieces to enrich the framework.

Thus, the novelty of the framework resides in the development of this workflow or procedure that encapsulates expert knowledge (agronomist or crop modeller) into crop models (crop simulators developed by a software engineer) for various applications. These different crop simulators will result from rules that explicitly state why a specific approach has been selected. Such a framework facilitates the comparison of different modelling approaches and enhances testing and comparing different approaches to particular processes, so that scientific understanding can advance. We are aware that there is rarely one unique modelling solution. However, we think that this framework can guide the user to define the best solution for a specific objective, given the available expert knowledge and the available data. Importantly, the expert knowledge is not static, but will evolve through the use of the framework, e.g. through comparisons of models composed of different modelling approaches for similar processes, or validation of results generated by different developed models using experimental data.

2.3. Software design to incorporate expert knowledge in the definition of crop simulators

To include new modelling approaches into the framework, a clear understanding of the software design is needed. The possibility to easily add a new modelling approach and combine it with already existing ones is based on software engineering techniques. Software engineering techniques such as object-oriented design have been applied for a long time in the industrial sector, but have only recently been introduced in the field of agricultural research (e.g. APES, Donatelli et al., 2010). In software engineering, the main purpose of a design pattern is to describe simple and elegant solutions to specific and recurrent problems (Gamma et al., 1995). Design patterns can also be used in the development of a crop growth and development modelling framework. In our particular case, we decided to use the “plug and play” architecture (Papajorgji, 2005).

The use of the plug and play architecture provides the capability of automatically creating new configurations, and is based on: (i) the strategy design pattern, which is a behavioural pattern: “defining a family of algorithms, encapsulating each one, and making them interchangeable” (Gamma et al., 1995), and (ii) the abstract factory design pattern which is a creational pattern: “providing an interface for creating families of related or interdependent objects without specifying the concrete classes”

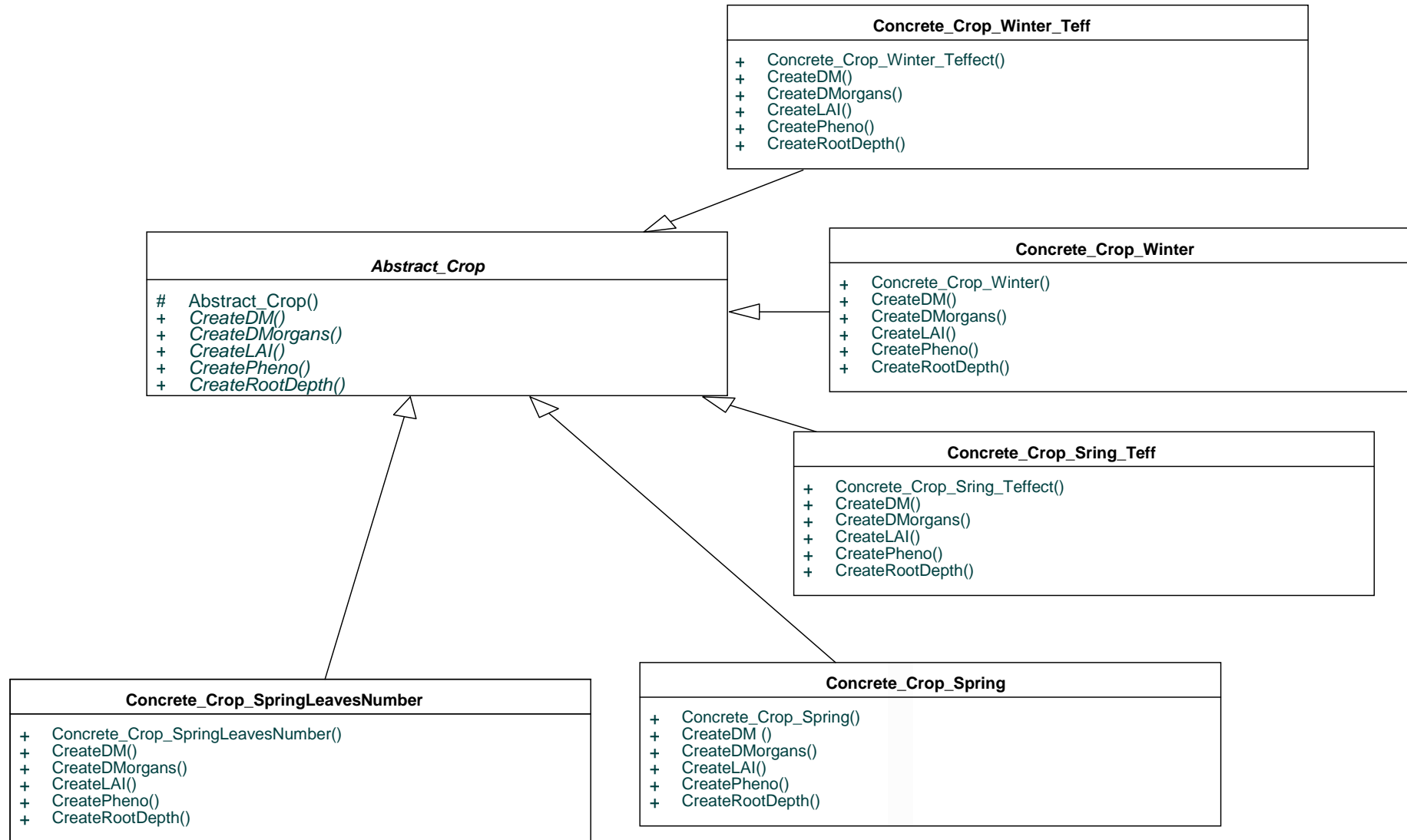
(Gamma et al., 1995). Strategy and abstract factory design patterns are combined in such a way that the strategy design pattern permits to set up the library of plant processes and the abstract factory design pattern allows to create the library of crop simulators (combination of these different processes). The strategy design pattern is also used to define the model granularity within APES components (Donatelli et al., 2010). A full description and concrete application of some of these components can be found in Donatelli et al. (2006a, 2006b, and 2010). Descriptions of other components (e.g. soil water, soil nitrogen and carbon) can be found on the APES website (<http://www.apesimulator.org/>).

The abstract factory defines the basic structure of the model. For instance, the abstract crop class includes various methods (e.g. `CreateDM()`, `CreatePheno()`, Figure 2.2a) that represent the crop growth and development processes. The different modelling approaches for each of these processes are encapsulated in various strategies (e.g. `BiomassProd_LUE` class or `BiomassProd_LUE_CS` class, Figure 2.2b). The strategy design pattern describes a family of mutually interchangeable modelling approaches. Each family of strategies defines the basic crop growth and development processes where exchangeability of modelling approaches is possible.

Furthermore, the abstract factory design pattern facilitates the creation of a new crop simulator in such a way that a new concrete factory can be added easily (e.g. `Concrete_Crop_Winter`, Figure 2.2a). The concrete factory is the reflection of the expert's perception of the system or/and the result of the different criteria checked by the user in agreement with his/her simulation objective. If none of the concrete factories available corresponds to the expert's choices, a new concrete factory can easily be created. Finally, the use of the abstract factory allows including a common approach in line with the generic models' principles.

Thus, through the use of the plug and play architecture (Papajorgij, 2005), we include in the framework an explicit link to the representation of the expert knowledge of the system. The creation of new concrete factories corresponds to the development of new crop model structures (i.e. crop simulator) and the use of the abstract factory provides flexibility in the choice of modelling approaches. For instance, if the user wants to simulate potential crop yield of winter wheat at field level, the choice of a winter crop in the GUI will direct him to pick the `Concrete_Crop_Winter` presented in Figure 2.2a, while if the user wants to simulate spring wheat, then the `Concrete_Crop_Spring` will be selected. In the same way, if the user wants to simulate crop yield of winter wheat for the whole of Europe, the `Concrete_Crop_Winter_Teffect` will be selected, as the range of temperatures within Europe might be an important characteristic to consider (Ewert et al., 1999).

a. The abstract factory design creating a basic changeable structure of the crop model.



b. The strategy design pattern applied to the CreateDM() corresponding to the dry matter process.

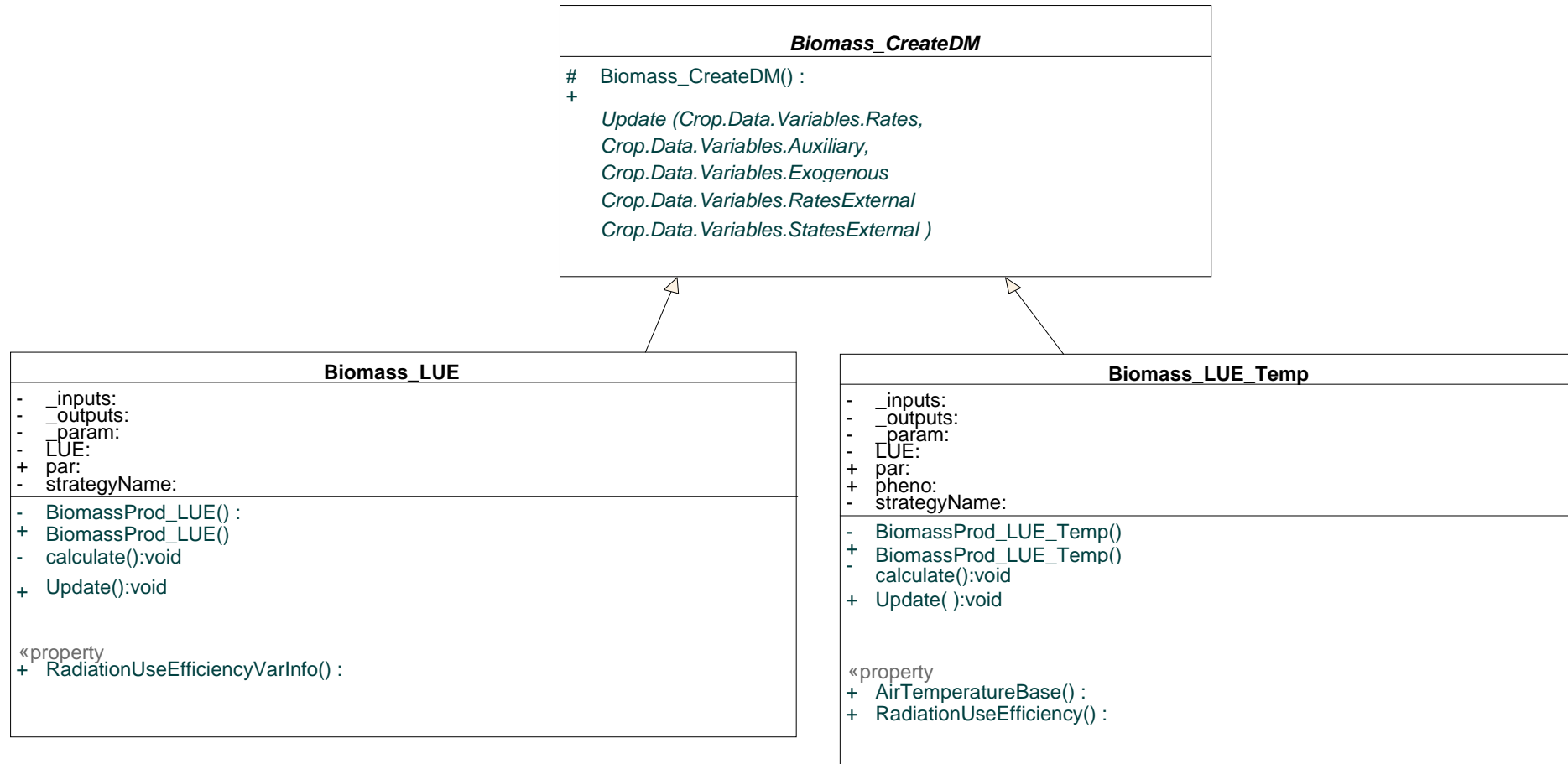


Figure 2.2. The main design patterns implemented to develop the crop modelling framework: a) the abstract factory design pattern and b) the strategy design pattern.

Not only does the application of this design prevent implementation of inconsistent combinations of objects, but it also provides flexibility to the framework. This flexibility lies in the possibility of creating new combinations of modelling approaches (i.e. crop simulators). However, if the user's overall vision of crop growth and development is totally different from the one formulated in the abstract factory, a change in the design is required. This is not a "shortcoming" of the pattern, but it rather points to the degree to which object-oriented systems are amenable to adding or removing abstractions. In that sense, the expert knowledge enters the framework at different levels. First, the abstract factory defines the key plant physiological processes according to crop physiologists' knowledge (which have been already pre-defined according to crop physiology, Section 2.1). Second, through the selection of criteria (GUI), the agronomist expert defines different modelling approaches (strategies) and ways to combine them (concrete factories). To illustrate the principles and the design described above, an example is presented to define different modelling approaches that can be used to represent the phenology of a crop and identify the appropriate approaches for a set of different research objectives.

3. Application of CROSPAL: example of crop phenology

As mentioned by Hammer et al. (2002), crop modelling is moving in two directions, one from cropping level to higher integration levels such as farms and regions and the second from crop level to the lower level of integration, such as the genetic level. We claim that the development of the proposed crop modelling framework supports both directions, as different modelling approaches can be included and considered for further application depending on the objective of the user of CROSPAL. We use the example of phenology to further illustrate this. Phenology is the sequence and rate of initiation of developmental events (i.e. phases) in the course of the life cycle of a crop, from germination of seeds to maturity of the plant and ripening of the seeds. The sequence of the different phases is invariable, but their timing is dependent on environmental conditions, such as temperature and day length. In evaluating crop growth models, phenology is the first process to be assessed because it is a driver of most of the other plant processes. It can be considered at 3 levels of detail (Table 2.2): (i) the crop level, based on the classical photo-thermal modelling approach, (ii) the gene network level, on the basis of a more detailed and mechanistic approach, and (iii) a more aggregated summary, based on the timing of flowering and ripening according to a simple regression on environmental factors. Each of these approaches has its specificities and targeted applications. The most suitable approach for a specific objective is selected on the basis of expert knowledge.

Table 2.2. Main characteristics of the modelling approaches studied to simulate phenology.

	Leaf appearance rate, and photo-thermal time approaches	Gene network model approach	Simple regression approach
Main objective	Provide a temporal framework for simulating crop growth processes defined by the appearance of leaf primordia and leaves	Understanding the consequence of modifying gene networks on the prediction of flowering time	Determine harvesting time at regional level
Type of approach	Mechanistic: Timing of each phenological stage based on qualitative morphological changes on the developing apex	Molecular pathway dynamics: mathematical formulation of gene network	Empirical relation: regression analysis between sowing date, latitude and flowering date or harvesting date
Main assumptions	It assumes a decline in the rate of leaf production as a function of temperature for leaves formed later in the crop cycle	For a specific cultivar, the alleles present in the network regulate aspects of temperature and photoperiod responses	Flowering and physiological maturity time at regional level is mostly dependent on sowing date and latitude
Main inputs	Daily temperature, day length	Genotype specific input	Latitude and sowing date
Crop characteristic/ Parameters	Thermal time of the different stages is cultivar dependent;	Development stages to define the photoperiod-sensitive period	Regression coefficient according the crop and the location
Main outputs	Development stages, date of heading and leaf numbers	Date of flowering, leaf numbers	Regional date of harvest

3.1. Definition of the different modelling approaches (different strategies)

The most commonly used approach to simulate crop development is based on the temperature sum (thermal time) above a base temperature. The base temperature represents the temperature below which development ceases. The development rate is defined as the inverse of the duration of a development phase. The approach assumes that within a range of temperatures, crop development rate increases proportionally with increasing temperature (Roberts and Summerfield, 1987). The crop enters the next development stage when the thermal time requirement for the current phase is satisfied (Van Keulen and Seligman, 1987). A similar approach is based on the rate of leaf appearance and final number of main stem leaves. Jamieson et al. (1995) demonstrated that: (i) leaf primordia appearance in wheat can be predicted on the basis of temperature alone, and (ii) final leaf number is controlled by the photoperiodic and vernalization responses of the crop. In this approach, the production rate of leaf primordia is considered to be independent of day length and sowing date (Miglietta, 1992).

A more mechanistic approach to simulate flowering time is based on QTL (Quantitative Trait Loci) analysis, determining the sensitivity of the crop to photoperiod. This approach enables to predict the effect of a modified gene network on flowering time (Van Oosterom et al., 2006). According to day length, different crop phenotypes will react differently. Integration of such an approach in a crop model enables to link information at the molecular level with the phenotypic performance of the crop. Analysis of the gene network gives a better insight in why a crop responds in a certain manner to different environmental conditions (e.g. day length).

A much simpler approach that can be used to predict flowering is based on empirical observations and simple regressions between flowering time and latitude (Hammer et al., 1996), or sowing time and the duration of the growth cycle. To derive such regressions for various crops, regional data for different crops in different regions are required. The main assumption underlying this approach is that sowing date is the most important factor in determining flowering and physiological maturity of a specific crop or variety in a region.

Each of these approaches can be implemented as a strategy according to the strategy design pattern. The choice of the strategy for a specific objective is done via the selection of the factory that corresponds to the criteria identified in the following section.

3.2. Selection of modelling approaches given an objective

Accurate prediction of crop development is essential for accurately simulating morphogenesis and yield components, and is also important in scheduling management practices and assessing the response of a specific cultivar to e.g. weather variability. All three approaches mentioned above have been applied in (different) crop models. One of these approaches will be most appropriate for a specific objective of the simulation. Consideration of expert knowledge will help to find the most relevant approach for each situation. One could imagine users of CROSPAL, having the following different modelling aims:

Aim 1: Explain the gene influence on photoperiod sensitivity of cultivars to flowering time;

Aim 2: Determine the right timing for management practices such as fertilizer application or disease control according to the development stage of the crop;

Aim 3: Study crop productivity and harvesting time in a long-term perspective (e.g. under different management practices) for different regions.

Table 2.3 summarizes the possible applications of phenology modelling approaches, and the selected model approaches according to the simulation objective or aim.

Table 2.3. Objectives, criteria and modelling approaches selected to simulate phenology.

Objective of simulation	Main criteria determining the choice	Approach chosen
Understanding the principle of development	Crop physiology: Need of a mechanistic, ontogenetic approach	Gene network model
Scheduling management practices	Management practices Fertilization: importance of the double ridge stage in wheat Pesticide application: number of leaves	Thermal time and Leaf appearance
Regional yield prediction	Scale: Simulation over a broad geographical range	Simple regression

Aim 1:

The main objective of this simulation is to increase understanding of the variation in the response of a crop or cultivar to environmental factors. More specifically, the user wants to predict the phenotypic consequences (e.g. time of flowering) for a crop with a new gene combination. The gene network model expresses the photoperiod sensitivity

of a crop in having either a recessive or dominant gene (Van Oosterom et al., 2006). Clearly, to achieve this aim, an eco-physiological QTL model should be incorporated into a crop simulation model. Such an approach will allow the user to test different gene combinations and determine how the crop responds to different environmental conditions.

Aim 2:

(a) The timing of fertilizer application should match the temporal variation in nutrient (e.g. nitrogen) demand. To accurately match supply with demand, it is important to identify the periods of high nitrogen requirements. Kirby et al. (1989) showed that in wheat the developing shoot apex is most sensitive to fertilizer application at the double ridge stage. This stage was shown to be best approximated by the thermal time approach (Miglietta and Porter, 1992) that allows explicit identification of the timing of each critical development stage such as the double ridge, terminal spikelet, ear emergence, etc. Therefore, from knowledge retrieved from the literature, we identified the thermal time approach as most appropriate in determining the timing of fertiliser application.

(b) The timing of plant disease control can often be linked to a specific development stage. However, for practical purposes it is often identified by a certain number of leaves (Bindi et al., 1995). Although the thermal time approach defines the temperature sum from one phenological stage to another, it does not take into account that thermal time between successive leaves may vary with cultivar, planting date and location. The leaf appearance rate approach does account for such differences. The approach determines the number of leaves and therefore the timing of a specific stage as characterized by a certain number of leaves, and holds over a wide range of sowing dates and latitudes. Thus, for determination of the timing of disease control measures, the leaf appearance rate approach is most appropriate.

Aim 3:

For studies aiming at estimating e.g. the long-term effects on harvest time and crop productivity of a farmer's adaptation of management practices to new policies, a simple approach that predicts the length of the crop cycle and of the grain filling period is needed. In long-term model predictions, it was demonstrated that empirical approaches, based on observations and simple regressions predicted yield as well or even better than more detailed agro-climatic crop models (Hammer et al. 1996). Indeed, due to poor knowledge of some input data, the use of a simple approach can lead to a good trade-off between accuracy and likely cost of application in a forecasting mode.

Therefore, if the main objective of the study is for example to define harvest time of

wheat over a long-time series at regional level (e.g. the SEAMLESS project, Van Ittersum et al., 2008), a very simple approach such as the regression model will be robust (and sufficient) for large-scale applications. The main criteria to decide on this regression model (between sowing date, latitude and harvesting date) are the scale (i.e. regional application) and the availability of data.

3.3. Workflow of the methodology and use of CROSPAL

Figure 2.3 shows the key steps of our methodology for the construction of a crop simulator using expert knowledge, combining knowledge from agronomists, crop modellers and software engineers. Following the methodology explained above (Section 2), first the crop modeller with the help of the agronomist identifies the different modelling approaches to describe phenology that should be included in the framework. Then, the software engineer translates these into strategies (step 2). In a third step, the agronomist, having his objective in mind, identifies the key criteria for the simulation of the crop system and selects them within the GUI. Finally, the software engineer creates the (different) concrete factories (if needed) according to the selection criteria considered in the GUI (step 4).

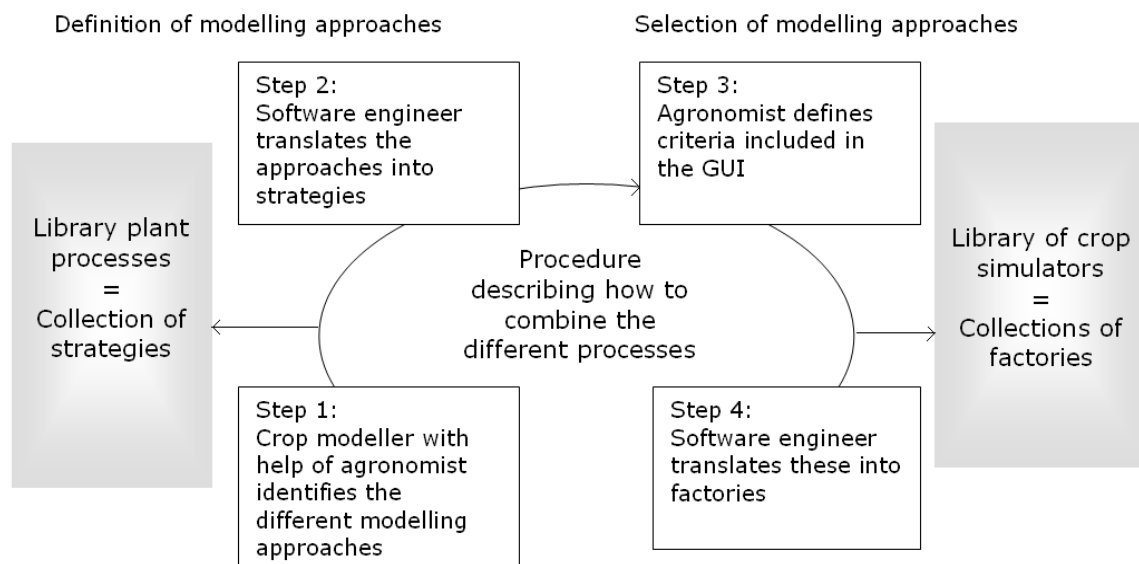


Figure 2.3. Key steps in the crop modelling framework to select the model structure on the basis of expert knowledge.

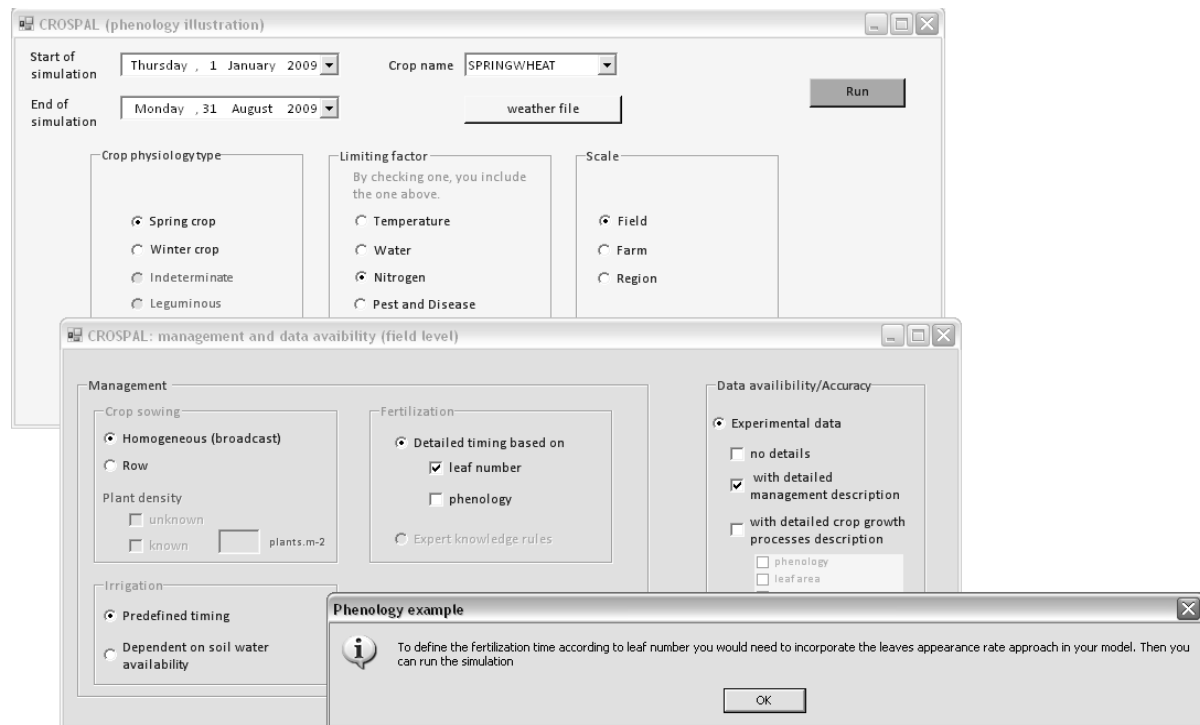


Figure 2.4. Main screens of the CROSPAL GUI used to schedule the timing of pesticide application for a wheat crop.

To illustrate the outcome of this methodology, we describe the case where the user of CROSPAL (e.g. an agronomist) wants to identify the best pesticide application time for a wheat crop (aim 2). The user of CROSPAL (e.g. an agronomist) sees the need to define some specific development stages, which are practically identified by a certain number of leaves (Bindi et al., 1995). Before the user of CROSPAL, an agronomist, starts to run CROSPAL, step 1 and 2 (inclusion of various modelling approaches within the software) have already been realized in collaboration with a crop modeller and a software engineer, as a starting point for the use of the CROSPAL. Then, running CROSPAL, the user will identify the main characteristics of his system following the GUI of CROSPAL (step 3, Figure 2.4): (i) spring wheat, (ii) growth in water and nitrogen limiting conditions, and (iii) field level. Resulting from this particular selection, a second window will appear where the user will have to select different criteria related to the management practices at field level (key criteria for this particular case). The user will define a homogenous field with a pre-defined irrigation schedule and the importance of the determination of leaf number, which will automatically require detailed experimental data (Figure 2.4). Following these various selections, the user interface of CROSPAL will indicate that a leaf appearance rate modelling approach is required and that it is not yet included in the framework.

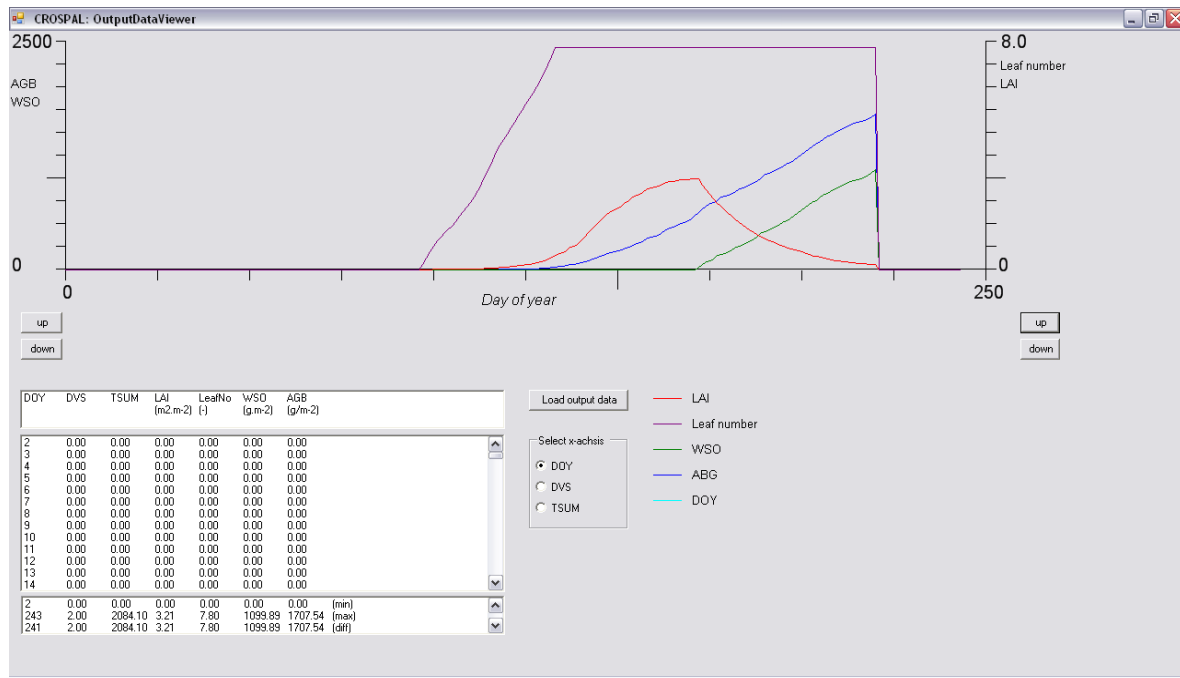


Figure 2.5. Main outputs of the crop simulator in CROSPAL resulting from the selection for simulation aiming at scheduling pesticide application for a wheat crop.

LAI: leaf area index; WSO: Weight Storage Organs; AGB: Above ground Biomass; DOY: Day Of the Year.

Finally, with support of a software engineer, the user will include this approach as a new strategy and define a new concrete factory (step 4). Figure 2.5 shows the results of the simulation after incorporation of the leaf appearance rate approach and the associated factory. This sequence of steps facilitates the use of a procedure in which the information will not be lost. Not only does such a methodology help to define various crop simulators, based on previous experiences/expert knowledge, but the proposed crop modelling framework also presents a clear procedure for the future modeller to create his/her own crop model structure from the existing library within the framework.

4. Discussion and conclusions

In this study, we described CROSPAL, a crop modelling framework for users to define different combinations of modelling approaches of crop growth and development processes to build a crop model for a specific objective. This framework focuses on the main challenge to adequately combine different crop modelling approaches of processes to a crop model with consistent and appropriate scales of process complexity (Voinov et al., 2004; Rastetter et al., 1992). CROSPAL relies on the explicit inclusion and utilization of expert knowledge through the definition of selection criteria in a GUI, the graphical interface between the software engineer (coding the strategies and factories) and the agronomist or crop modeller (defining the main basic crop processes and how to combine them).

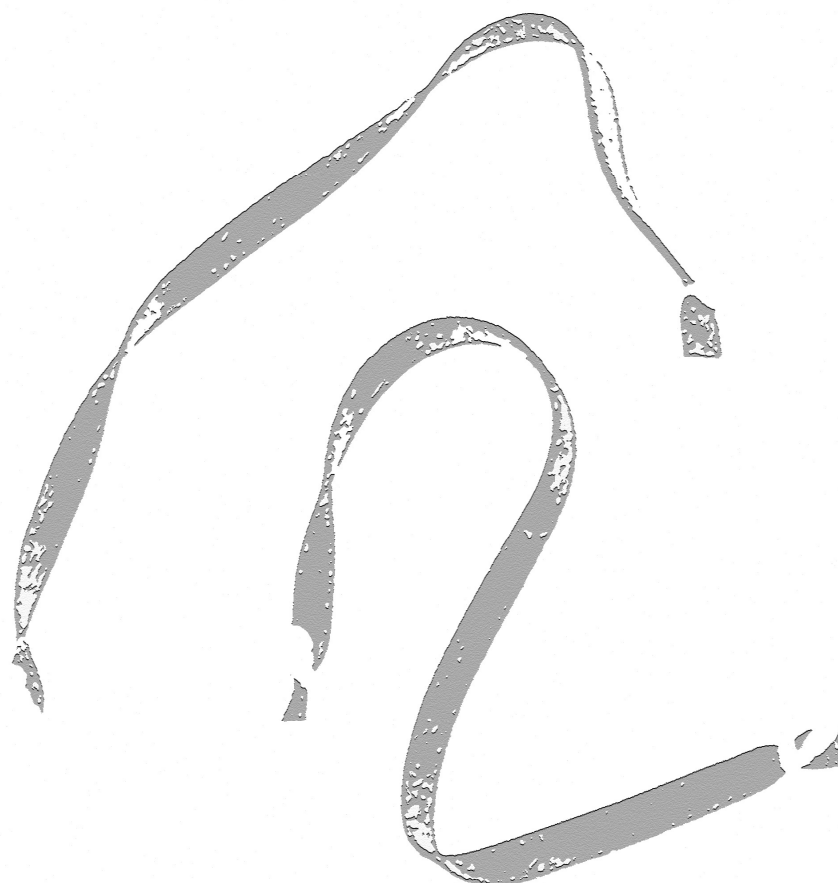
As Rahman et al. (2004) mentioned, a modelling framework should comprise different modelling approaches and should guide the user in running a simulation. CROSPAL contains a library of modules simulating various crop growth and development processes under various assumptions and guides the user in picking and assembling them to create a new library of crop simulators. The proposed design builds on the work from Wang et al. (2002), but explicitly integrates expert knowledge in the selection of the relevant structure of the crop model depending on its objective. The user (e.g. an agronomist) will, in general, have knowledge on the crop physiology and on the impact of agronomic practices on crop growth and development. He/she will specify the simulation objective on the basis of the five criteria identified in the framework: (i) crop physiology, (ii) limiting factors, (iii) scales, (iv) management practices, and (v) the desired accuracy of simulation. The expert may not necessarily need a strong background in modelling or software engineering, as the proposed framework will provide a tool to translate their systems knowledge into modelling options.

As mentioned by Holzworth et al. (2010), abstraction is “a key technique to enable [...] separation of logic”. In CROSPAL, the assemblage of the abstract factory and the strategy design patterns provides the framework with a high flexibility, and structures the combination of the modelling approaches, the logic of the crop simulator. The flexibility resides in the possibility to create a new strategy (ease of extensibility) or/and a new concrete factory (ease of composition), while the abstract factory provides a consistent structure to combine processes. By using this type of design, we create for the user the possibility for an explicit link between the different crop modelling approaches included in the framework (the strategies) and the representation of the system by expert knowledge (the factories). However, as mentioned above, the definition of the abstract factories can be difficult to maintain if the user’s overall vision of crop growth and development is totally different from the one presently represented. Indeed, we defined the abstract factory according to the key plant physiological processes that we pre-defined according to crop physiology (Section 2.1). If the user disagrees on this delineation, a change in the design will be required. Moreover, although CROSPAL is extensible, future development of the software will require a strong and efficient interaction among agronomists, crop modellers and software engineers (Chapter 5). Such work relies on effective collaboration among these different disciplines and therefore illustrates the importance of integrative science.

Finally, we illustrated the applicability of these principles with an example on crop phenology. We have selected phenology, as this process is central in many crop models and the first process to consider when testing simulation models of crop production (e.g. Ewert et al., 1999). Once this process is accurately modelled, other processes such as resource capture and allocation can be studied. The main focus of this study was to describe a procedure for combining different modelling approaches based on expert knowledge. Therefore, we choose to illustrate our approach with an example where the knowledge on the processes (phenology) is not limiting. However, our approach should also allow addressing issues that deserve special attention (heuristic goals) in the field of crop modelling such as resource competition models (including interaction with soil modules). Such models will require to include other strategies regarding for example water dynamics (e.g. water uptake, water demand) and to explicitly define criteria (e.g. soil water holding capacity, crop drought resistance) that will result in new factories (Chapter 3-4-5) and additions to the GUI. Additional effort will be needed to further develop and extent CROSPAL for this and other applications. Importantly, not only should CROSPAL represent a way to capitalize what is known in the field of crop modelling through pre-establishing crop simulators but also serve as a learning tool for the researcher, crop modeller and/or agronomist to test new hypotheses on plant growth and development processes.

CHAPTER 3

Systems analysis: the modelling beam of the framework*



* This chapter is under review as:

Adam, M., Belhouchette, H., Corbeels, M., Ewert, F., Perrin, A., Casellas E., Celette, F., Wery, J., 2010. Systematic evaluation to support model selection in a modular crop modelling framework: an application for simulating crop response to nitrogen supply. *Agricultural Systems*.

Parts of this chapter were presented as :

Belhouchette, H., Adam, M., Casellas, E., Celette, F., Corbeels, M. & Wery, J., 2009. Performances of two crop models in various conditions: The importance of underlying assumptions. In M.K. van Ittersum, J. Wolf & H.H. van Laar (Eds.), *Proceedings of the Conference on Integrated Assessment of Agriculture and Sustainable Development: Setting the Agenda for Science and Policy (AgSAP 2009)*, Egmond aan Zee, the Netherlands, 10 - 12 March, 2009 (pp. 196-197).

Adam, M., Ewert, F., 2009. How to deal with modularity and flexibility in crop modeling? Oral presentation on the German Society of Agronomy. Halle, Germany. September 1-3, 2009.

Abstract

Crop models require different structures for different applications. Modular crop modelling frameworks, such as the recently developed agricultural production and externalities simulator (APES), support the change of model structure. However, the assembly of different modules to create a model may not always result in the best model structure. We developed and tested a protocol for a systematic selection and evaluation of a crop growth model structure. The novelty of the presented approach relies on a throughout analysis of the different modelling approaches (modules) and on how to assemble them to create new modelling solutions (i.e. model). We use a case study to demonstrate that we can explicitly express and test the different assumptions behind the choice of a specific modelling approach. Our case study refers to the simulation of crop growth in response to nitrogen management and the importance of an accurate simulation of the nitrogen uptake. Applying the proposed protocol, we identify the need to improve the initially selected nitrogen mineralisation module. We conclude that the protocol is suitable to provide guidance for systematic testing of different crop processes modelled and that the crop modelling framework can be extended by including various modelling approaches to improve a modelling solution, instead of developing a plethora of individual crop growth models. The use of the protocol highlights the importance of the documentation of the modelling process and of the clarification of the uncertainty associated.

Keywords: model structure, modules, uncertainty, selection, protocol, nitrogen.

1. Introduction

Much progress has been made in developing mathematical models that simulate the development and growth of crops under various conditions of weather, soil and management; widely known models are e.g. CropSyst (Stöckle et al., 2003), APSIM (McCown et al., 1996), DSSAT (Jones et al., 2003), EPIC (Williams et al., 1989), WOFOST (Supit et al., 1994). Most of these models are designed to operate as one comprehensive entity with a specific structure and detail for simulating crop growth and soil nutrient (often nitrogen) cycling processes (e.g. EPIC, Sharpley and Williams, 1990, STICS, Brisson et al., 2003). Nevertheless, the degree of detail included in simulation models should correspond to the specific research question addressed (Passouria, 1996). Models should be as simple as the nature of their objectives allow, not be overloaded with unnecessary details, and have minimum data requirements (Sinclair and Seligman, 1996). Therefore, it should be possible to construct *ad hoc* crop growth models customized to specific simulation problems. This asks for modular crop modelling frameworks. For few of the existing crop growth simulation models, such modularity is achieved by providing a set of modules with different degree of complexity for a specific crop or soil process, often meant to adapt the model to the available input data (e.g. CropSyst provides the user with a choice of sub-models to predict evapotranspiration, Stöckle et al., 1994).

In SEAMLESS-IF (a computerized framework to assess and compare *ex-ante* alternative agricultural technologies and agricultural and environmental policy options, Van Ittersum et al., 2008), the crop modelling framework APES (Agricultural Production and Externalities Simulator, Donatelli et al., 2010) has been designed to be flexible and modular. APES consists of various modules for simulating crop growth and development, soil carbon, nitrogen and water dynamics. Its design facilitates the adjustment of model structure depending on the objective of the simulation, the data availability and the type of cropping system (annual crops as well as grasslands, vineyards and agro-forestry under a range of soil/weather conditions and management practices). Such a crop modelling framework offers the flexibility to potentially different ways of combining modules to one effective simulation model (herein referred to as modelling solution, MS). A key challenge in using modular and flexible frameworks such as APES is to find an MS which is the “best fit” for a specific application. It requires an approach based on systems analysis to select the best MSs, or if necessary, to create a new, better MS.

The main objectives of the present paper are (i) to describe a protocol for the selection and evaluation of a modelling solution and (ii) to demonstrate and test its use with an application to a case study, in which we assessed the impact of nitrogen management on crop production and externalities in Flevoland, the Netherlands.

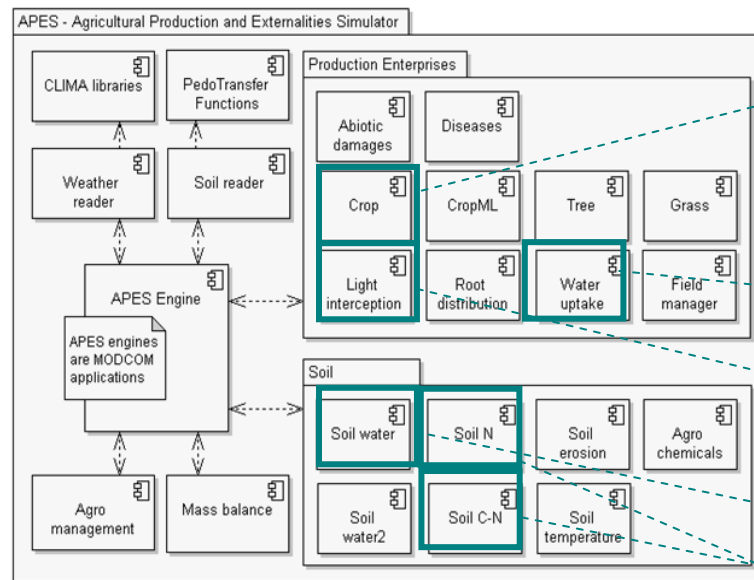
2. Material and Methods

2.1. Description of the crop modelling framework APES

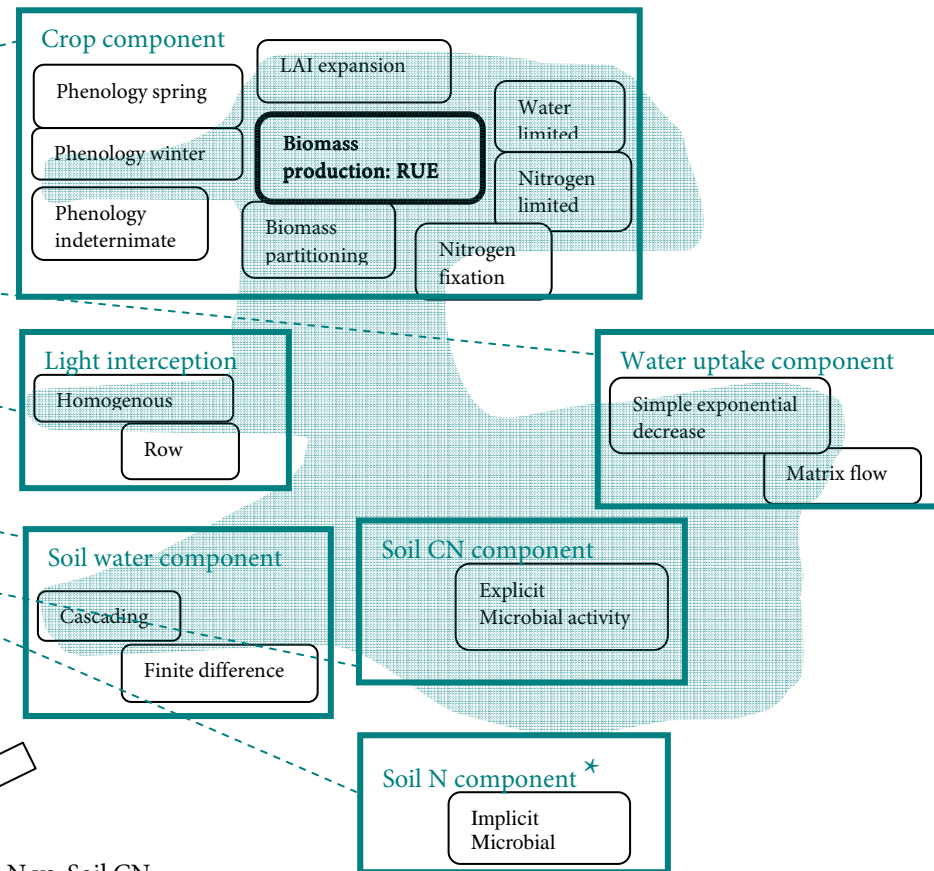
APES is a modelling framework based on the concept of modularity to improve flexibility in handling the diversity of cropping systems and simulation applications. While the diversity of cropping systems can be the result of various factors including: (i) biophysical conditions (soil, weather), (ii) type of crop, land use system or agro-management (grassland, cereal, legumes, perennial crops, agro-forestry), and (iii) type of production (fodder, grain, tuber), the range of simulation applications can be characterised by e.g. (iv) type of crop performance indicators and environmental externalities evaluated (nitrogen leaching, erosion, soil carbon) and, (v) data availability for model evaluation and simulation (experimental data, expert knowledge, statistical data). The modularity of APES is illustrated by the possibility to include different components and modules representing crop and soil processes.

Figure 3.1 depicts the different levels of granularity included within APES and how a modelling solution (MS) corresponding to the simulation objective can be defined from components and modules. A component is a piece of software representing crop and/or soil process that is used to compose a cropping system model (e.g. crop, light interception, water uptake, soil water, soil N, soil C-N, components highlighted in Figure 3.1a.). A component can comprise various modules. A module is a specific conceptualization of a crop or soil process implemented within a component (e.g. radiation use efficiency, RUE , for biomass production within the crop component, Figure 3.1b.). Flexibility in APES is given through the ability of easily combining different components and modules to create diverse MSs (3.1b and 3.1c), i.e. a specific combination of different modules and components depending on the objective of the simulation (e.g. the simulation of crop yield and nitrate leaching in a specific type of land use system with specific management practices), and on data availability (type of data and their quality).

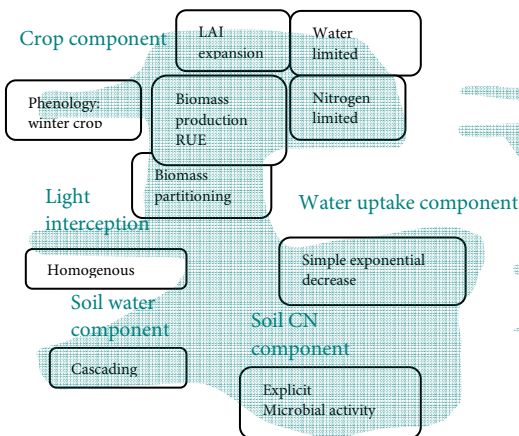
a. select **COMPONENTS** from APES



b. select **MODULES** from the selected component to create a **MODELLING SOLUTION (MS)**, with a special focus on crop growth



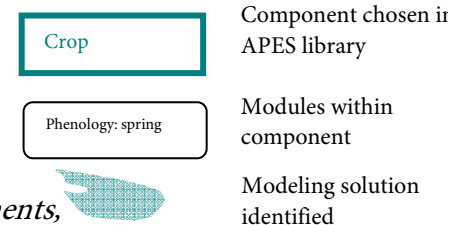
c. add the MS to the **MS LIBRARY**



* Soil N vs. Soil CN

Sometimes a component can be identified as a module if:

1. no alternative module is included and
2. there is a second component for the same process



39 Figure 3.1. Representation of APES modularity and flexibility with different levels of granularity: (a) the components, software units used for composition, (b) the modules, conceptualization of a process within a component, (c) the modelling solutions (MS), combination of different modules depending on the objective of the simulation.

Currently, APES integrates various modules representing crop phenology (determinate vs. indeterminate), crop canopy dynamics (i.e. leaf area index -*LAI*- expansion and senescence), dry matter production and partitioning with the effects of stress factors, root growth, all including in the crop component; the water dynamics in the soil (cascade approach vs. Richards equation, non-linear partial differential equations to represent water flow in the soil), part of the soil water component; and nitrogen dynamics in the soil included in two different soil nitrogen components (one with microbial activity regulating crop residue decomposition and one with an implicit representation of the microbial activity, based on first-order decomposition rates). Table 3.1 gives an explicit list of the current components and modules available in APES (Donatelli et al., 2010). At the same time, APES has been designed to allow further extension of this list of modules, when necessary and inclusion of other component if required (e.g. crop disease component, Salinari et al., 2008).

2.2. Protocol for selection and evaluation of a modelling solution, based on systems analysis

Because the APES crop modelling framework offers different ways of combining components and modules to one MS, we developed a protocol for identifying the best MS for a specific application. The protocol follows three main steps: (i) MS selection, (ii) MS calibration, and (iii) MS evaluation and improvement. Each of these steps includes four aspects (facets) of analysis: the criteria of selection and evaluation, the problem to define, the available data, and the selection and evaluation of model components (Figure 3.2). The protocol leads to two main outcomes: (i) an MS selected and evaluated for a specific application, and (ii) an associated uncertainty matrix (Walker et al., 2003) that identifies the potential “unknown” related to the MS for the targeted application.

The uncertainty matrix distinguishes different types and sources of uncertainties in order to facilitate uncertainty classification. The matrix is divided in four main parts: (i) the contextual part, referring to the uncertainty related to the understanding of the system under study, (ii) the input/data, analysing the uncertainty related to the dataset, (iii) the parameters, associated with the calibration process of the MS and the uncertainty related to the value and meaning of parameters, and finally (iv) the model structure highlighting the parts of the model where knowledge is not yet complete.

Table 3.1. Components and modules available in APES.

APES: libraries of components and modules	
Components	Modules Available
Light interception component	<ul style="list-style-type: none"> • Homogenous • Pronk (Pronk et al., 2003)
Crop component	<p>Phenology modules</p> <ul style="list-style-type: none"> • Thermal time • Photothermal time • Photovernaltime • Indeterminate phenology <p>Leaf area expansion module</p> <ul style="list-style-type: none"> • Biomass accumulation dependent (Spitters and Schapendonk, 1990) <p>Dry matter production module</p> <ul style="list-style-type: none"> • Radiation use efficiency (Monteith, 1977) <p>Partitioning/allocation module</p> <ul style="list-style-type: none"> • Predetermined allocation (Van Keulen and Seligman, 1987) <p>Water dynamics module</p> <ul style="list-style-type: none"> • Water stress index moderated with a drought tolerance parameter <p>Nitrogen dynamics modules</p> <p>Nitrogen stress based on the NNI approach (Lemaire , 1997; Shibu et al., 2010)</p> <ul style="list-style-type: none"> • Nitrogen stress on RUE (Green , 1987) • Nitrogen stress on RUE and LAI (Vos et al., 2005)
Water uptake component	<ul style="list-style-type: none"> • Water uptake is defined by using parameters such as root conductance and leaf potential
Soil water component	<p>Two water dynamics modules</p> <ul style="list-style-type: none"> • Simple cascade approach • Richard's equation approach
Nitrogen component (i.e. SoilN)	<ul style="list-style-type: none"> • Soil nitrogen available: nitrogen transformation process is driven only by water and temperature (Johnsson et al., 1987)
Soil CN component	<ul style="list-style-type: none"> • Soil nitrogen available. The role of soil micro-organisms is represented in a mechanistic way through the mineralization-immobilization turnover processes during organic matter decomposition (Corbeels et al., 2005)

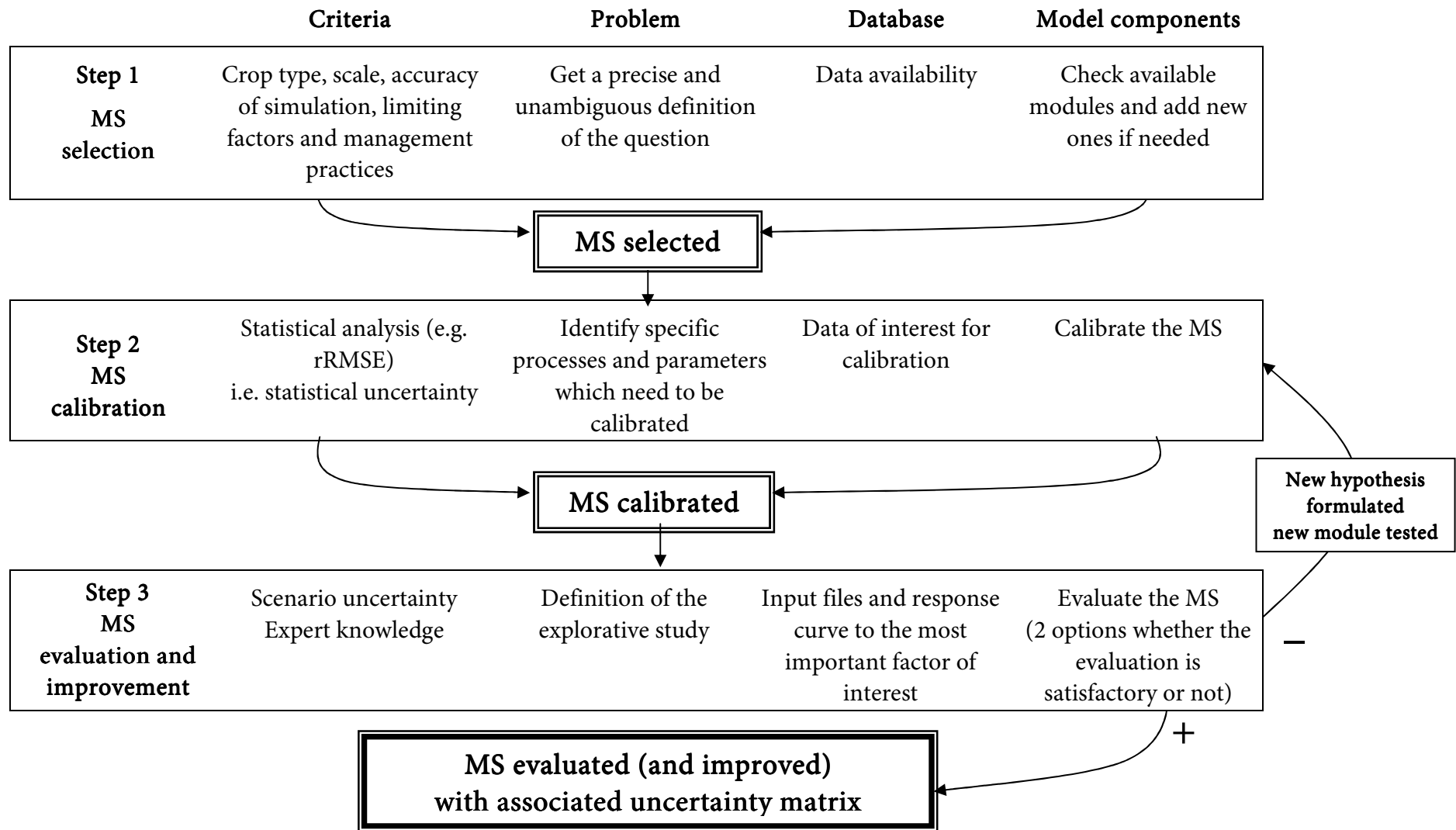


Figure 3.2. Protocol for a systematic evaluation of a modelling solution (MS) within the crop modelling framework. This protocol follows three major steps: (1) MS selection, (2) MS calibration, (3) MS evaluation and improvement.

2.2.1. Step 1- MS selection

In the first step of the protocol, we explicitly describe the rationale and objective of the model application according to basic criteria derived from the crop modelling expertise. We define five main criteria to guide the selection of modules: (i) crop type, (ii) spatial and temporal scale, (iii) required accuracy of simulation, (iv) growth limiting factors, and (v) management practices (Adam et al., 2010). These criteria consider genotype x environment x management interactions (Yin et al., 2004), including effects of technological changes (Ewert, et al., 2005), scale and data availability. These criteria are used to assess the suitability of the available components and modules in APES to the objective of the simulation and to the available data. The clear definition of the modelling goal leads to a clear description of the cropping system: the boundaries, the components and its environment (external factors). Such description contextualizes the model application by (i) getting a precise and unambiguous definition of the question to be addressed with the model and (ii) identifying the relevant external factors and agro-management options to be simulated with the model. The contextualization helps to accurately delineate which processes are relevant for the specific application and the uncertainty associated with the definition of the system under study. Then, in agreement with the application objective, we identify the data needed and clarify whether they are available for the specific application. The identification of the available data, allows defining the limitations associated with these data for model calibration and input variables. This step of the protocol leads to the selection of the appropriate model components and modules to assemble an MS and to a clear identification of the contextual uncertainty (i.e. problem framing) of the application, as a result of the explicit definition of conditions and circumstances of the system studied. If some modules important for the application are not available, we need to create and add them to the framework, before to proceed with the following step.

2.2.2. Step 2- MS calibration

The second step is essential to gain confidence in the selected MS and to improve its applicability for the specific study. Specific statistical criteria used for the evaluation of the MS are the root mean square error (*RMSE*) or the relative mean absolute error (*RMAE*), as an indication of the importance of the model error and its distribution, or the coefficient of correlation (*R*), an indication of the degree to which the observed and simulated variables trends are associated. An important first action in this step is the proper identification of the processes and parameters that need to be calibrated. The processes considered depend on the objective of the application, the selected

components and modules, and the available data. For instance, if no leaf area index (*LAI*) data is available, a detailed calibration of *LAI* dynamics is impossible. Through this step in the protocol, we gain information about the importance of the parameters for the specific application and the reliability of the parameter values due to the data available (or not) for calibration. Thus, the uncertainty related to the available data and parameters of the models is clearly identified.

2.2.3. Step 3- MS evaluation and improvement

The main objective of this third step of the protocol is to evaluate the MS by applying the calibrated MS to a wider range of conditions and to analyze how the calibrated MS behaves under this broader range of agronomic conditions. The decision criteria to accept the MS as a fair representation of the crop's behaviour is when variables (indicators of the system behaviour) reproduce the observed response of the crop to the factor of interest. These observed responses can also be expectations based on expert knowledge or results in the literature, but are ideally independent data. To create a wide range of soil-weather-agro-management situations, we test the sensitivity of the model to the different factors of interest for the specific application (e.g. temperature, nitrogen input). Then, the MS is analysed looking at the main model output variables representing the overall performance of the cropping system such as the state variables representative of key crop processes, i.e. above-ground biomass, grain yield, leaf area index, water and nitrogen uptake, and stress indices. If the selected MS does not reproduce the expected behaviour potential hypotheses on the reason of this mismatch are formulated and other module(s), derived from other crop models or expert knowledge, is(are) selected or included in the framework. The modularity of the framework enables this easy plug-and-play of components for further testing, until we establish the “right” MS for the specific application.

2.3. Case study

2.3.1. Objective

It is anticipated that due to high prices fluctuations and environmental constrains (i.e. compliance to the Nitrate Directive of the European Commission, EC, 2001) farmers will have to adapt their management to more efficient and sustainable cropping practices. In this case study, we look more explicitly at nitrogen management and its impact on crop yield and externalities (nitrate leaching) in the Flevoland region of the Netherlands. Applying nitrogen beyond the crop needs leads to undesirable nitrogen leaching and unnecessary costs for the farmers. Efficient nitrogen management should prevent from such drawbacks and favour yield productivity.

2.3.2. Data sources

We collected data from the literature for the three main crops present in Flevoland (winter wheat: Groot and Verberne, 1991; maize: Horsch, 2000; and potatoes: Van der Schoot et al., 2002). In these datasets, crop growth and development was monitored in detail to capture the crop dynamics within the growing season. We tested our selected MS for various management practices: sensitivity to different level of nitrogen application representative for the Flevoland region.

2.3.3. Data analyses

To calibrate the selected MS, we used the relative mean absolute error ($RMAE^1$), to measure how close the outputs of the model (S_i) were to the observed data (O_i). For the sensitivity analysis (i.e. crop response to different nitrogen management), we analysed the model outputs with the expected behaviour (i.e. expert knowledge). More specifically, we used the De Wit 3-quadrant (De Wit, 1992) to analyse the crop yield responses to different nitrogen management practices, looking at nitrogen capture and conversion efficiency by the crop as a function of soil available nitrogen. We also look at nitrate leaching, to assess the potential undesirable losses of nitrogen of the system.

3. Results

Table 3.2 summarizes the application of the protocol to the case study assessing the impact of nitrogen management on crop production and nitrate leaching in Flevoland. Table 3.3 represents the associated uncertainty matrix.

3.1 Step 1: MS selection

Criteria of selection: The 5 criteria for MS selection are given in Table 3.2 (as mentioned in section 2.2.1). Following these criteria, the problem definition and data available are analysed to guide the model components selection.

Problem definition: The main objective is to assess the impact of different nitrogen management practices on agricultural productivity and environmental externalities (i.e. nitrate leaching) in one European region (i.e. Flevoland, the Netherlands). The main crops grown in Flevoland are winter wheat, potatoes and fodder maize. In terms of the required accuracy of the simulation, we are interested in the model behaviour, analysing the yield response to different levels of N supply, rather than “exact prediction” of yield. We assume that the crops do not experience any significant water limitation during the growing cycle, mostly because of high rainfall in Flevoland (tested with simulation, see the uncertainty matrix, Table 3.3).

¹ $RMAE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|S_{ij} - O_{ij}|}{S_{ij}} \right)$ with i (number of observations) = 1,...,n and j the different variables considered

Table 3.2. Application of the protocol for a targeted application with the objective to assess management practices on crop production and externalities in Flevoland.

CASE STUDY	Criteria	Problem	Database	Model components
Step 1 MS* selection (5 criteria)	1. Crop type	Winter wheat, maize and potatoes	Winter wheat	Photo-vernal-thermal phenology module
	2. Scale	Region	Field experiment	Leaf area expansion is biomass accumulation dependent
	3. Accuracy of the simulation	Pattern	Crop growth processes data. No soil data.	Radiation use efficiency (<i>RUE</i>) Partitioning is with predetermined allocation table
	4. Limiting factor	No water limitation		Simple cascade approach and no water stress
	5. Management practices	Nitrogen management	Different nitrogen treatments Poor information on soil characteristics	SoilN and Nitrogen stress on <i>RUE</i>
Step 2 MS Calibration	Statistical evaluation <i>RMAE</i> < 0.2	Calibration phenology, <i>LAI</i> dynamics, <i>RUE</i> and nitrogen content in the crop	Selection of the experimental dataset with highest nitrogen application	Calibration first on potential production conditions and then on nitrogen limiting conditions (results in Table 3.4)
Step 3 MS evaluation and improvement	Typical Liebig curve De Wit 3 quadrant	Test the sensitivity of yield to a gradient of nitrogen Analysis of the nitrogen use efficiency	Sensitivity: Different nitrogen application 2 growing seasons Data to evaluate : winter wheat with different nitrogen treatments	Evaluation/ improvement MS : too low nitrogen uptake and too high nitrogen leaching <div style="border: 1px dashed black; padding: 5px; margin-top: 10px;">New module to test: Inclusion of a more mechanistic representation of the microbial activity in the soil nitrogen dynamics (SoilCN)</div>

* MS: modelling solution

Table 3.3. Uncertainty matrix associated with the MS to simulate crop response to nitrogen application in Flevoland, NL (adapted from Walker et al., 2003).

Source of uncertainty	Nature System or knowledge-based <i>The “known known”</i>	Range Statistical / Scenario <i>The “unknown (to be) known”</i>	Recognized ignorance Mostly due to data availability <i>The “known unknown”</i>
Contextual: boundaries and definitions			
Flevoland, the Netherlands	Weather data from Swifertbant Soil information representative of the experimental location	While running models at water limited conditions, no water stress was observed	We considered that there is no water limitation as we are in Flevoland (NL)
Winter wheat, maize and potatoes	Most extensive dataset on winter wheat (Groot and Verberne, 1991)	Not enough data on maize and potatoes	–
Nitrogen (N) management	Three treatments: 80-140-240 kg ha ⁻¹	–	–
Input/data uncertainties			
Field experiments	Experimental trials in PAGV 1982-83 carried out for simulation studies	“Several parameters were measured in a rather crude manner or derived from general soil data bases” (Groot and Verberne, 1991p.349)	Growing conditions are not always optimal even in experimental trials (despite conscious monitoring)
Parameters			
Calibration	Winter wheat growth and development under potential growth conditions	Phenology, leaf area index and dry matter dynamics, <i>RUE</i> and crop N dynamics	Unclear distinction between the different crop organs for N content; no information on N leaching
Explorative study	Effect of nitrogen stress Nitrogen uptake	We tested nitrogen stress effect on LAI and <i>RUE</i> : no differences were observed	We assume that the N stress affects only the <i>RUE</i>
Model Structure			
Biomass production and LAI	Yield, Nitrogen uptake	Yield, nitrogen content in leaves and storages organs	Assumption: good simulation of crop N uptake = a good simulation of the soil N dynamics
Crop nitrogen dynamics modules	N effect on <i>RUE</i> ; N uptake	N stress effect on <i>LAI</i> and <i>RUE</i> : no differences observed	We assume that the N stress affects only the <i>RUE</i>
Soil modules	Microbial uptake	N use efficiency: N uptake	No precise information of soil N available for the crop and soil microbial activity

Database: We collected data from the literature for the three main crops. The most extensive dataset was for winter wheat (Groot and Verberne, 1991). For this particular dataset, phenological stages, *LAI*, dry matter in green leaves, grains and total above ground biomass, as well as the nitrogen content of the different crop organs (leaves, stems, grain and roots) were monitored at weekly intervals. Field experiments were conducted during two growing seasons with different levels of nitrogen application. Daily weather data were obtained from the meteorological station at Swifterbank, 15 km from the experimental farm. The datasets for the other crops only included a couple of harvesting dates (Van der Schoot et al., 2002) or were located in the Netherlands but not precisely in Flevoland region (Horsch, 2000).

Model components: Considering the data available and the objective of the application with the criteria defined previously, we defined the different modules/components that can be selected.

(1) The most extensive dataset being on winter wheat, we decided to focus our study on this crop. It requires a phenology module that includes a vernalization process. We selected a photo-vernal-thermal approach (Stöckle et al., 1994) available in APES, to represent the phenology of winter crop (Table 3.1).

(2; 3) Although the scale for this study was the region, we used a field experimental dataset. This was acceptable as in Flevoland the agro-environmental conditions (climate and soil) are considered homogeneous. Further, we looked for data over the growing season to capture the model behaviour in response to farm management actions, asking for a model with a rather low input data demand, but running at a daily-time step. Therefore, we looked at rather simple representation of the crop growth processes, also in agreement with our criteria of simulation accuracy: looking at model's behaviour in response to one limiting factor, rather than quantitative prediction. The following main modules (responding to these criteria) were selected within the crop component of APES (Table 3.1): (i) leaf area development, with the approach of the leaf area expansion dependent on biomass allocation (Spitters and Schapendonk, 1990); (ii) biomass production, with the radiation use efficiency approach (Sinclair and Muchow, 1999); and (iii) biomass allocation with predetermined allocation coefficients (Van Keulen and Seligman, 1987).

(4) As water was not considered as a major limiting factor (Table 3.3), we used the simple cascade approach for simulating water movement in soil (i.e. infiltrating water is passed on layer by layer down the soil profile as upper layers are refilled to field capacity). Coupling of this approach with the soil nitrogen uptake and mineralisation allows us to assess nitrate leaching. No gaseous losses are simulated.

(5) Concerning the nitrogen dynamics, i.e. soil nitrogen mineralisation, we considered a rather simple soil nitrogen dynamics module (soilN, Table 3.1), not requiring detailed calibration, as the data availability in terms of the soil nitrogen dynamics were rather scarce (i.e. the data collected were more related to crop growth processes than to soil nitrogen dynamics: *LAI*, dry matter and nitrogen in the crop). The soilN module is based on the main nitrogen transformation processes driven only by water and temperature (Johnsson et al., 1987). The soil microbial biomass is implicitly represented as part of the litter and manure pools. It was coupled with the crop growth modules via a nitrogen stress effect (nitrogen nutrition index, NNI approach, Lemaire, 1997). Looking at winter wheat, the nitrogen stress was assumed to mostly affect the dry matter accumulation via *RUE*, with no effect on *LAI* (Green, 1987, Shibu et al., 2010). A list of the different modules selected for the initial MS is provided in Table 3.2, step 1.

Uncertainty matrix: From this first step of the MS selection, we explicitly stated that no water limitation occurred in Flevoland, mostly because of the high rainfall in the study area. We tested this hypothesis before to continue further. While running the model with rainfed conditions, we found that the crop did not experience any water stress, as we expected (contextual part of the uncertainty matrix, Table 3.3). Further, we identified rather poor data availability in terms of the soil database, leading to the selection of soil components with rather simple water and nitrogen soil dynamics approaches.

3.2. Step 2-MS calibration

Problem: While selecting the MS, we considered that the most important processes to calibrate are (i) the phenology, (ii) the biomass production and partitioning (iii) leaf area of the crop, and (iv) crop nitrogen dynamics (assuming that a good simulation of crop nitrogen uptake is a pre-requisite for a good estimation of nitrate leaching). Therefore, the parameters to calibrate should relate to these four main processes in the crop component (Table 3.4) and more specifically to (i) the temperature sum corresponding to the various development phases, (ii) the specific leaf area and the relative growth rate of *LAI* during its exponential growth (i.e. juvenile phase), (iii) the radiation use efficiency and reallocation of biomass from leaves and stems to the storage organs during the grain filling period, and (iv) nitrogen concentration in the different crop organs (i.e. leaves, roots, stems and storage organs).

Table 3.4. Main crop growth module parameters for the APES-MS after calibration and the associated RMAE for the main variables calibrated.

Modules calibrated within the crop component	Parameters (after calibration)			Calibration evaluation	
	Description	Values	Unit	Variables calibrated	RMAE
Phenology	Air temperature (Base)	0	°C	Estimated from observation and weather information	
	Air temperature (Sum Emergence)	100	°C d		
	Air temperature (Sum Anthesis)	600	°C d		
	Air temperature (Sum Maturity)	730	°Cd		
	Maximum number of days for vernalization	32	days		
Leaf area development	Relative growth rate of <i>LAI</i> (exponential phase)	0.006	(°Cd) ⁻¹	Green leaf area index	0.12
	Specific Leaf Area	0.021	m ² g ⁻¹		
Biomass production and partitioning	Radiation Use Efficiency (<i>RUE</i>)	3	g MJ ⁻¹	Dry matter in green leaves	0.19
	Fraction of dry matter reallocated from leaves to the grains (storage organs)	0.4	-	Dry matter in above ground biomass	0.11
	Fraction of dry matter reallocated from stems to the grains (storage organs)	0.2	-	Dry matter in grains	0.23
Nitrogen dynamics	N Max concentration in storage organs	0.025	g N gDM ⁻¹	N content in leaves	0.12
	N Max concentration in leaves at DVS 0.5	0.05	g N gDM ⁻¹		
	Fraction Max N concentration in root from N Concentration in leaves	0.37	-	N content in grains	0.20
	Fraction Max N concentration in stems from N Concentration in leaves	0.40	-	N content in whole plant	0.28
	Translocation N (Time Coefficient)	6	days		

Database: To calibrate the MS, we used the experimental data with the highest nitrogen application, to simulate potential growth conditions (i.e. no nitrogen stress on crop growth). We performed the calibration in a step-wise manner. We first estimated the phenological parameters from the observations and recorded weather data. Then, we calibrated parameters related to leaf area index, biomass production and allocation and nitrogen dynamics, respectively. The calibration was done manually by trial-and-error method, comparing the model results with the observed results. We evaluated the outputs with the *RMAE*, considering a value below 0.2 to be acceptable (Mayer and Butler 1993).

Calibration of MS: Table 3.4 gives the parameters calibrated and the calculated *RMAE* for the main variables used for calibration. For phenology, the accumulated temperature sum for physiological maturity has been calculated to assess the harvesting time of the crop; the accumulated temperature sum at anthesis corresponds to the observed date of maximum leaf biomass and beginning of storage organ growth. The vernalization requirement of winter wheat has been adjusted to observe a start of an “effective” crop growth at the beginning of February. With respect to the leaf area index dynamic, the default value of relative growth rate of leaf area index during the juvenile phase [$0.009\ (^{\circ}\text{Cd})^{-1}$] is based on Van Keulen and Seligman (1987) who included a wide range of data from different wheat cultivars (i.e. both spring and winter wheat) in their analysis. Winter wheat requires a lower value, corresponding to the range observed by Van Delden et al. (2001). The specific leaf area was estimated from the observed data, to avoid compensation error with the calibration of the radiation use efficiency (*RUE*). The range of variation for *RUE* is small, as it is more a crop specific than a variety specific parameter. Consequently, once total biomass has been calibrated (through *RUE*), calibration of reallocation parameters enables to adjust the biomass partitioning between the different organs at the end of the growing season. Finally, for the nitrogen dynamics, the calibration was carried out on nitrogen content of the various crop organs and the overall crop uptake at the end of the growing season. Figure 3.3 shows the model outputs after calibration compared to observed data for dry matter and nitrogen content in the various crop organs and for the leaf area index.

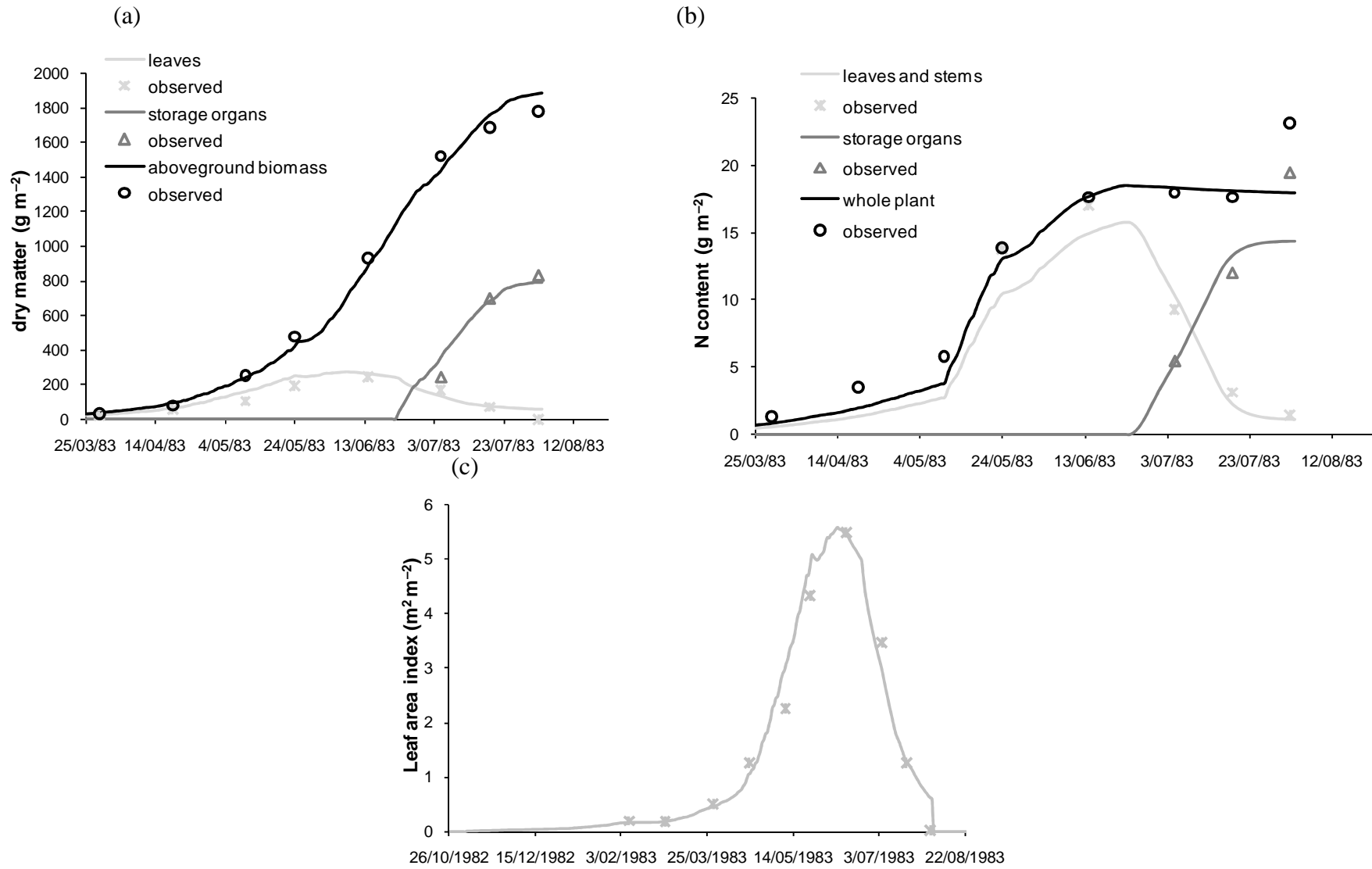


Figure 3.3. Observed and simulated after calibration (a) dry matter, and (b) nitrogen content in various crop organs and (c) leaf area index during the 1982-83 growing season.

Most of the *RMAE* have a value under 0.2 (Table 3.4), proof of a satisfactory calibration of the MS. The higher value of the *RMAE* (but still acceptable) for the dry matter in grain can be related to slower dry matter accumulation at the beginning of grain filling as the main difference between the observed and simulated values were observed during this period rather than at the end of the growing cycle (Figure 3.3a). The higher value for the total nitrogen content in the crop can be related to an underestimation of the nitrogen content in the roots for which we did not have any observed value. Also leaves, sheaths and stems were separated in the observed data, while in the model sheaths are not explicitly simulated. Also, as shown in Figure 3.3b, the nitrogen content in grains increased suddenly at the end of the growing season, while the model did not simulate such “unexpected” change.

Uncertainty matrix: The main uncertainty for this step remains in the nitrogen dynamics. First as mentioned earlier, the distinction between the different crop organs in the experiments was done differently than in the model (Table 3.3). Further, the calibration was mostly carried out on crop processes and no detailed evaluation of soil nitrogen dynamics was carried out. We assumed that if the crop nitrogen uptake was properly simulated, the soil nitrogen dynamics were also simulated in a reasonable manner (Table 3.3).

3.3. Step 3-MS evaluation and improvement

Problem: The aim of the third step of our methodology is to test whether our selected and calibrated MS responds correctly under a wide range of levels of nitrogen fertilization. To test the sensitivity of yield to a gradient of nitrogen, we set up an explorative study where the input data vary only in terms of nitrogen input.

Database: We ran simulations with weather data and soil information from the dataset of Groot and Verberne (1991). Their dataset comprises detailed crop measurement for two growing seasons (1982-83 and 1983-84) with various nitrogen treatments, which we used for evaluation of the MS. However, with respect to the soil dynamics, they clearly stated that soil parameters were not always collected in sufficient detail and could only be derived from a general database (see input/data uncertainties in Table 3.3). When necessary, we used results from other simulation exercises (e.g. nitrate leaching) conducted previously with the same dataset (De Willigen 1991). We varied the nitrogen applications from 0 kg N ha⁻¹ to the maximum of 240 kg N ha⁻¹, with 80, 140 and 180 kg N ha⁻¹ as intermediate treatments, in agreement with the experiments (Groot and Verberne, 1991).

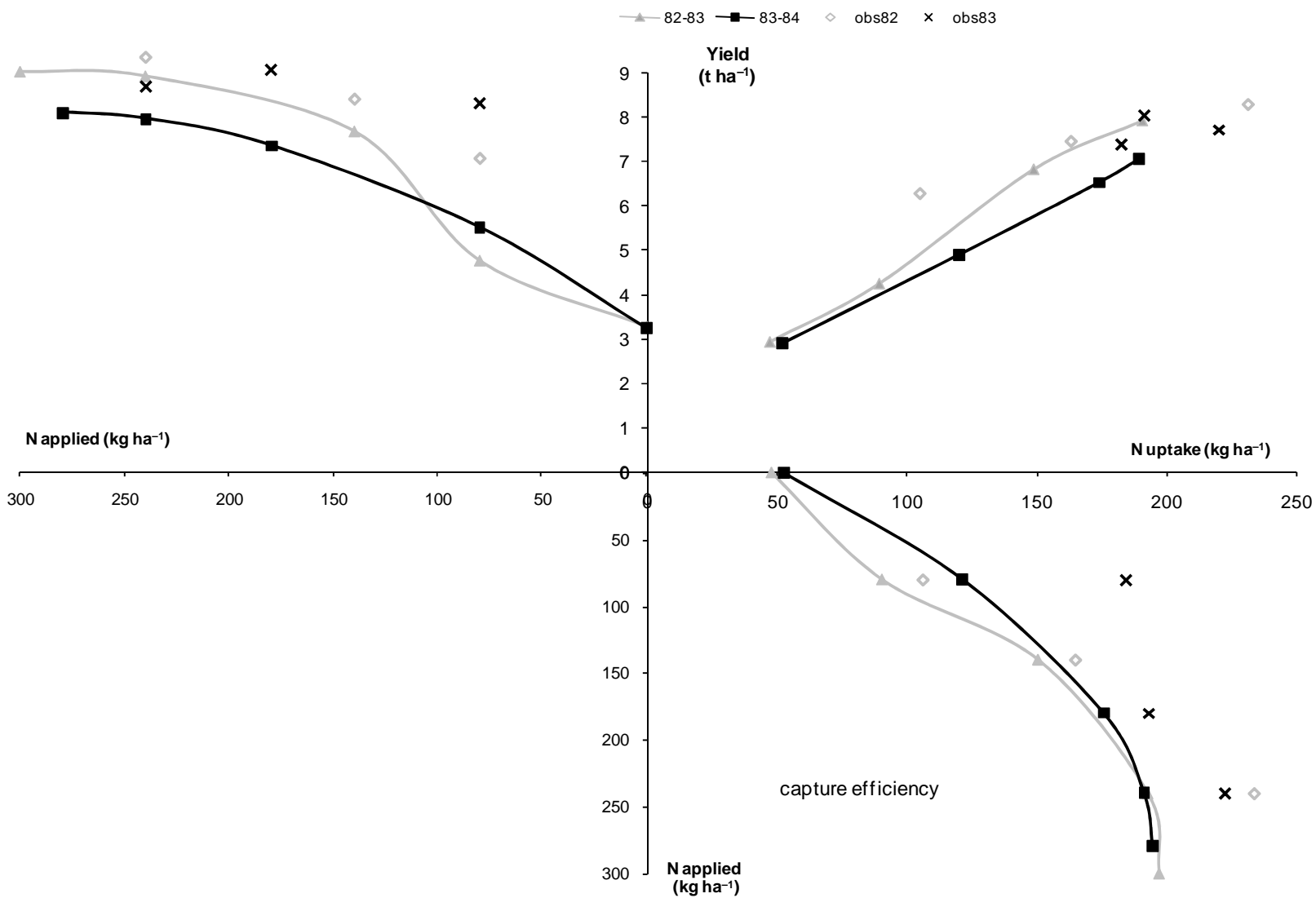


Figure 3.4. The 3 quadrant diagram to analyze the crop response to nitrogen application (after De Wit, 1992).

MS evaluation: Figure 3.4 represents the crop response to the nitrogen gradient. We used the De Wit 3-quadrant representation (1992) to analyse the nitrogen use efficiency by the crop. One of the quadrant presents grain yield against nitrogen application (i.e. fertilizer response curve, representative of the agronomic efficiency, AE); the second, yield against nitrogen uptake (yield-uptake curve, representative of the physiological nitrogen use efficiency, NUE); and the third quadrant relates nitrogen application to nitrogen uptake (application –uptake curve, representative of the apparent nitrogen recovery, ANR).

Simulated grain yield varies from 3t ha⁻¹ for the no nitrogen application to 8t ha⁻¹ for the highest nitrogen application (i.e. 240 kg N ha⁻¹). The fertilizer response curve does not follow the typical response curve determined by Liebig (Quadrant II, De Wit, 1992). It appears that although the curve reaches a plateau for an optimal nitrogen supply, the simulated nitrogen uptake (Quadrant III) is lower than the observed nitrogen uptake (for all doses of the nitrogen application) and nitrate leaching appears to be high (even when no nitrogen fertilization occurred, Table 3.5). From previous simulation exercises (De Willigen, 1991), nitrate leaching was considered negligible.

The soilN module selected considered the crop nitrogen demand according to a logistic curve (Greenwood et al., 1974). During a crop growth period, the main mechanisms behind the nitrogen uptake are diffusion and root absorption and the nitrogen uptake is only limited when the soil nitrogen concentration is low. The nitrogen available in the soil is the result of the net mineralisation in which implicitly it is considered that the soil microbes absorb a fraction of the available nitrogen. However, according to Leffelaar et al. (2007, pers. comm.), the proportion of nitrogen uptake by the microbes is far from negligible and should not be under-estimated. Therefore, we concluded that the fraction (already) absorbed by the microbes in the soilN module might be too low for the soils in Flevoland, which have high organic matter content. Underestimation of the potential microbial nitrogen uptake could lead to an overestimation of nitrogen leaching (and as a result an underestimation of crop nitrogen uptake). Therefore, we decided to include a new soil nitrogen dynamics module (SoilCN, Table 3.1) which includes an explicit simulation of the microbial activity to test the above hypothesis. After calibration (results not shown), we estimated results (grain yield and nitrogen uptake for the calibration run) similar to the ones observed for the different nitrogen applications (Table 3.5). With respect to the explorative study, with this new module, nitrate leaching was negligible and although the crop nitrogen uptake tended to be overestimated, the simulated yield agreed better with the observed ($rRMSE$ = 0.1 compared to 0.2 for the initial MS).

Table 3.5: Main outcomes from the sensitivity analysis.

Growing season	Nitrogen application = scenario $kg N ha^{-1}$	Grain yield $t ha^{-1}$			Crop nitrogen uptake $kg N ha^{-1}$			Nitrate leaching*** $kg N ha^{-1}$		Nitrate leaching + uptake $kg N ha^{-1}$	
		observed	initial improved		observed	initial improved		initial MS	improved MS	initial MS	improved MS
			MS*	MS		MS	MS				
1982-83	0		2.91	4.43		47	99	48	1	95	100
	80	6.25	4.22	6.99	105	89	160	47	1	136	161
	140	7.44	6.80	8.12	163	149	204	47	1	196	206
	240**	8.28	7.89	8.34	231	191	227	47	1	237	228
1982-83	0		2.87	3.94		52	89	82	7	134	95
	80	7.36	4.88	5.87	182	120	159	82	7	202	166
	180	8.03	6.51	7.61	191	174	210	82	7	255	217
	240	7.7	7.05	7.78	220	189	215	82	7	271	222

*MS: Modelling Solution; the initial MS results from the first modules selection while the improved MS results from the selection of a new module after evaluation of the initial MS, in this specific the soilCN.

** This experiment (fertilization of $240 kg N ha^{-1}$, year 1982-83) was used for calibration. Calibration was carried out on the dynamics of leaf area index, dry matter accumulation and nitrogen content in the different crop organs, and not specifically on accumulated variables (i.e. total grain yield or total nitrogen uptake).

*** No observed data; assessment made from previous simulations (De Willigen, 1991).

Uncertainty matrix: The lack of data on soil nitrogen dynamics does not allow us to fully understand why we underestimated the crop nitrogen uptake and overestimated the nitrogen leaching while using the soilN module. We attributed this to an underestimation of the fraction of mineral soil nitrogen absorbed by the microbes. Use of a new soilCN module confirms such hypothesis. However, data collection on nitrogen mineralisation might help to better understand the phenomena simulated. Further, as there is no consensus in the literature on the nitrogen effect on crop growth (Olesen et al., 2002, Shibu et al., 2010), we also tested the module with nitrogen effect on *RUE* and *LAI*. No difference in the results was observed (see model structure in the uncertainty matrix, Table 3.3).

4. Discussion

4.1. Need of a protocol to select an MS in a modular crop modelling framework

The systematic approach for selecting and evaluating a modelling solution (MS) presented in this study is in line with the classical method of model building presented by Rabbinge and De Wit (1989), which includes 3 major phases: (i) conception, (ii) comprehension and (iii) evaluation and application.

In practice, the potential user of the framework must in the first place clearly define the purpose of the model application, before selecting and using a crop simulation model. Until recently, there were two options when using a model: (i) reuse an already existing model without introducing significant modifications or (ii) build a new model ‘*ad hoc*’ (Passioura, 1996). With the adoption of new software engineering techniques by the crop modelling community (Wang et al, 2000; Donatelli et al., 2002; Papajorgji and Pardalos, 2006; Adam et al., 2010), a third way of developing and using crop simulation models came into the picture: (iii) construct models from building blocks available in a crop modelling framework. The flexibility and modularity of APES, due to its component-oriented design, enables an easy technical assembly of these different modules in the same platform. But associated methodology should be provided to facilitate the selection and evaluation of models (i.e. modelling solution) (Ahuja and Howell., 2002). Not only should the crop modelling framework be an “implementation framework” (Van Evert et al., 2005) considering each module as a black box, but also a “modelling framework” (Van Evert et al., 2005) to formalize how to use these various modules.

Thus, our protocol should be seen as an important addition to the crop modelling framework to facilitate the decision on which MS is the most suitable for a specific application. It improves the process of creating a sound MS within APES. The protocol has two main advantages in the process of model use. First, it helps to contextualize the model application (Villa, 2007) by (i) getting a precise and unambiguous problem definition and by (ii) identifying the relevant modules to consider for the construction of the model. Second, the protocol enables (i) to evaluate the selected MS for the given application and (ii) to change/alter it (by substituting modules) if the selected structure does not capture the essential processes that are relevant for the particular application. The use of the protocol helps to test alternative hypotheses and evaluate the consequences (i.e. models results) of using other modelling approaches/solutions.

4.2. Need of documentation of the modelling process

The protocol promotes the compliance to the good modelling practices (Van Oijen, 2002) with the clear definition of the underlying principles for each module (Scholten, 2008) and the explicit definition of an associated uncertainty matrix to the MS, to reveal the potential “unknown” of the system (Refsgaard et al., 2007). The use of the protocol re-enforces the need of a good definition of the level of detail needed in a crop process-based model and should prevent from an “over-tuning” of parameters (Van Oijen, 2002), while emphasizing the potential uncertainty in the model structure. The protocol facilitates a systematic selection of different modules to re-assemble properly and create a new modelling solution. However, it does not preclude from a misuse of the crop modelling framework. Specific expertise is still needed to define the right component for the right simulation objective. For instance, in this study, while we tested two modules of soil nitrogen dynamics included in the framework (APES), we could notice differences in models’ behaviour (Table 3.5).

However, we were unable to fully understand why we underestimated the crop nitrogen uptake and overestimated the nitrogen leaching with one of the modules. We attributed it to a rather low implicit microbial activity contribution in the nitrogen uptake (Corbeels et al. 2005), especially for the soil under study (high organic matter content) and tested such hypothesis with a new module. However, such statement has to be taken with caution, and it re-enforces the need for good module documentation (Donatelli et al., 2002, Bellochi et al. 2010), when modules are included in the framework. The lack of documentation and a lack of expertise by the framework user in one or the other aspect of the modelling study might lead to a misuse of the different modules.

Table 3.6. Synthesis of the information given in the uncertainty matrix associated with our improved MS: qualitative assessment (1: rather certain; 0: low uncertainty; -1: high uncertainty).

Source of uncertainty	Nature <i>The “known known”</i>	Range <i>The “unknown (to be) known”</i>	Recognized ignorance <i>The “known unknown”</i>	Level of uncertainty (sum per source of uncertainty)
Contextual: boundaries and definitions				
Flevoland, the Netherlands	1	0	0	3
Winter wheat, maize and potatoes	1	0	0	
Nitrogen management	1	0	0	
Input/data uncertainties				
Field experiments	1	-1	0	0
Parameters				
Calibration	1	1	-1	0
Explorative study	0	0	-1	
Model Structure				
Biomass production and LAI	1	0	-1	-3
Crop nitrogen dynamics	0	0	-1	
Soil modules	-1	0	-1	
Level of uncertainty (sum per type of uncertainty)	5	0	-5	—

We described the use of the uncertainty matrix (Walker et al., 2003) as a way to tackle such issue by (i) explicitly stating where the lack of understanding was, and (ii) classifying the various sources of errors. Subsequently, we synthesized the information given in the uncertainty matrix, in a qualitative way, to comprehend better where the main uncertainties in our case study came from (Table 3.6). From this synthesis, we could identify (i) the lack of data on soil nitrogen dynamics as one of the main obstacles, (ii) as well as the incomplete understanding of the nitrogen mineralisation related to the nitrogen crop uptake. As a result, the main uncertainty in our MS remains in the model structure, but it could be explained by our “recognized ignorance” of some biophysical phenomena (N mineralisation and uptake), mostly related with a lack of data for further analysis. Such analysis clearly shows how our protocol helps to report on the decisions made during the modelling process. It also clearly highlights the potential misuse of the crop modelling framework when there is lack of documentation on the underlying assumptions of each module.

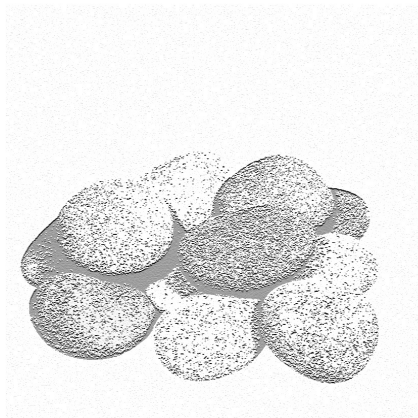
4.3. Concluding remarks

Our approach builds on the classical method of model building (Rabbinge and De Wit, 1989), mostly through its integration within a modular crop modelling framework. Model building should not be seen anymore as a linear process but rather as an iterative process where different hypotheses can be more easily tested due to the progress of software techniques. Use of the protocol presented in this study should guide the user in the selection of modules to assemble an MS. Indeed, modular crop modelling framework in combination with the protocol allows (i) to easily change the structure of the MS according to the simulation objective and data availability, and (ii) to define various MS with an explicit identification of the associated uncertainty, each corresponding to a specific simulation objective.

Future users of modular crop modelling frameworks will most likely be confronted with the same issues (i.e. flexibility) as the ones we faced in this study and they will most likely redefine new MS or/and include new modules in the framework. Thus, reuse of the protocol for new modelling applications will (i) further test the modules within APES, (ii) further refine the criteria for a more general applicability of the guidance for modules selection (Chapter 4), and (iii) further enrich the current library of modelling solutions of the framework (Figure 3.1), gathering different modelling approaches (Table 3.1) in one tool rather than in a plethora of models.

CHAPTER 4

Criteria of selection:
data availability and scale*



* This chapter is accepted for publication as:

Adam, M., Van Bussel, L.G.J., Leffelaar P.A., Van Keulen H., Ewert, F., 2010. Effects of modelling detail on simulated potential crop yields under a wide range of climatic conditions. *Ecological Modelling*, Accepted.

Abstract

Crop simulation models are widely applied at large scale for climate change impact assessments or integrated assessments. However, often a mismatch exists between data availability and the level of detail in the model used. Good modelling practice dictates to keep models as simple as possible, but enough detail should be incorporated to capture the major processes that determine the system's behaviour. The objective of this study is to investigate the level of detail incorporated in process-based crop growth models, usually developed and tested at the field scale that captures the effect of climatic variability on simulated yields for larger scale applications. We conducted a multi-site analysis and identified the impact of the effect of temperature and radiation on radiation use efficiency (*RUE*) on a daily basis, as well as on a seasonal basis. Further, we found that particular attention should be given to the choice of the light interception approach for large scale application of crop models. Two different *LAI* dynamics approaches (i) gave significant differences in simulated yields irrespective of the characteristics of the location and (ii) explained best the differences in the yield sensitivity to climatic variability. After clarifying the assumptions underlying the parameter representing the onset of senescence in both *LAI* dynamics approaches, the higher yields simulated by the summarized approach were attributed to a misrepresentation of leaf senescence. We concluded that a better understanding of leaf senescence is still needed, particularly to represent the onset of senescence in crop models.

Keywords: crop growth model, climatic variability, photosynthesis, radiation use efficiency, leaf area index, model complexity, leaf senescence.

1. Introduction

A key rule in good modelling practice is that the choice of a model depends on the question asked (Van Waveren et al., 1999). In crop modelling, a large variety of models has been developed since the 1960s, with each new model addressing a specific objective. Crop models have initially been developed for application at the field scale. Application of these models at larger scales such as for climate change impact assessments (Ewert, 2004a; Leemans, 1997) or integrated assessment studies (Van Ittersum et al., 2008) has become a common practice. However, for these applications, the required scale and objective of a crop growth model may go beyond the scope of the original model. Hence, the reuse of a model without any adaptation might lead to inaccuracies in model outputs, caused by (i) a misrepresentation of processes in the model, (ii) incorrect input data (unsuitable temporal or spatial resolution, or inaccurate measurements), including parameter values (Scholten, 2008), or (iii) a misinterpretation of the system, as the importance of effects relevant at lower levels may decrease at higher levels, while other factors, often not considered in the original crop model may become more important (Ewert, 2004a).

Challinor et al. (2004) identified the need for process-based crop growth models to capture the impact of climatic variability on crop yields over large areas. One of the challenges to apply a model for simulating crop growth and development at higher aggregation levels (e.g. Europe, Therond et al., 2010) is to ensure that the model appropriately addresses the response of crops to the temperature and radiation gradients found in such a heterogeneous environment. The model must reproduce the behaviour of the system under a wide range of conditions, representing the spatial variability, and therefore the model used should be robust. Bondeau et al. (2007) also mention the use of process-based crop models at the global scale to improve the representation of feedbacks between crop physiology and climate. A process-based model integrates descriptions of the underlying processes of the cropping system to explain its behaviour at the higher system level (Van Oijen, 2002), and usually includes at least two essential processes for crop growth, namely light interception by the leaf area and light utilization to produce biomass (Ewert, 2004b). In various crop growth models ((Ritchie and Otter, 1985); Spitters and Schapendonk, 1990; Spitters, 1990; Jamieson et al., 1998; (Stöckle et al., 2003); Bondeau et al., 2007) we found that (i) leaf area index (LAI , m^2 leaf area m^{-2} ground area) dynamics and (ii) biomass production are modelled with different mechanistic detail, usually depending on the main objective of the model developed.

A distinction can be made between explanatory, i.e. detailed mechanistic approaches with a high level of modelling detail, and descriptive, i.e. summarized approaches with a lower level of modelling detail (Penning de Vries et al., 1989). Detailed models have a high explanatory power, containing most of the elements and interactions that characterise a system, but they are resource-intensive (e.g. in terms of input data and simulation time). Summarized (also called summary) models are easier to handle (e.g. less parameters are needed and the models are simpler to interpret), but are generally more descriptive, reflecting little of the mechanisms explaining the behaviour of the system, often containing simplified representations of complex processes. Selection of the appropriate level of detail for each process to include in a crop growth model is often seen as a critical step in model development (Brooks and Tobias, 1996). It is a common rule to keep the model as simple as possible given the objective, but enough detail should be incorporated to capture the major processes that determine the system's behaviour (De Wit, 1968). It is also acknowledged that an optimum situation exists in terms of explanatory capacity of a model and the number of processes considered (Leffelaar, 1990; Passioura, 1996; Tittonell, 2008).

The objective of this study is to investigate the effect of the level of detail incorporated in a process-based crop growth model, usually developed and tested at the field scale, to simulate yields at larger spatial scale, i.e. regional level and higher. Particular focus is on *LAI* dynamics (representing the light interception approaches) and biomass production (representing the light utilization approaches) under potential growing conditions (Goudriaan and Van Laar, 1994; Van Ittersum and Rabbinge, 1997). We do not aim to develop the “best” large scale crop growth model, but the results of this study should improve the understanding of the relative importance of the different approaches to simulate potential crop yields at larger scales, especially in response to spatial differences in terms of radiation and temperature regimes.

2. Materials and Methods

We compared models with different modelling detail of the key growth processes of light interception and light utilisation to simulate crop yields in response to spatial variability in climate. The analysis follows two main steps:

1. Test of the models (i.e. modelling approaches) against measured data to ensure that all approaches are able to reproduce observed growth under field conditions for a range of climatic conditions. We selected experiments from different locations across the world that provided measurements of biomass and *LAI* and the associated weather data;

2. Systematic comparison of the behaviour of the models under a wide range of climatic conditions (following a climate gradient across Europe) to investigate the sensitivity of yield simulations to the use of the different approaches (i.e. different levels of detail) for the two growth processes examined.

2.1. Descriptions of the modelling approaches

We defined a detailed approach as one that describes a feature (e.g. crop growth) in terms of processes occurring at underlying hierarchical scale (also referred to as a more mechanistic approach). Incorporation of a description of photosynthesis according to Farquhar et al. (1980) can be considered as a detailed approach. In contrast, a summarized approach is seen as one that includes existing knowledge from underlying processes expressed in a simple relationship that describes the main responses of those processes to biotic and abiotic factors (also referred to as a descriptive approach, no explanation provided). Sinclair and Muchow (1999) identified application of the radiation use efficiency concept for simulating biomass production from intercepted radiation as an example of such an approach.

In this study, both, light interception by the leaf area and light utilization producing biomass were studied in a summarized and a detailed approach. Table 4.1 includes the key equations of the approaches studied. The associated parameter values can be found in Table 4.2.

Table 4.1. Key equations of the approaches applied for (a) LAI dynamics and (b) biomass production.

(a)

Equations for the detailed LAI dynamics approach

Juvenile phase ($LAI < 0.75$ and development stage < 0.16):

$$(1) \frac{dLAI_{\text{growth}}}{dt} = LAI \times R_g \times T_{\text{eff}}$$

with: $T_{\text{eff}} = \max(0, [T_{\text{average}} - T_{\text{base}}])$

T_{eff} , effective physiological temperature in °C and T_{average} , average daily temperature in °C

Following the juvenile phase:

$$(2) \frac{dLAI_{\text{growth}}}{dt} = \frac{dW_{\text{leaf}}}{dt} \times S_{\text{LA}}$$

After anthesis ($fT_{\text{sum sen}}$) or in case of self-shading (LAI_{critical}):

$$(3) \frac{dLAI_{\text{sen}}}{dt} = -R_d \times LAI$$

with: $R_d = \max(R_{d\text{-ag}}, R_{d\text{-sh}})$, see Appendix A.

Equations for the summarized LAI dynamics approach

Before anthesis:

$$(4) fLAI_{\text{max}} = \frac{fT_{\text{sum}}}{fT_{\text{sum}} + e^{(l_1 - l_2 \times fT_{\text{sum}})}}$$

with:

$$fT_{\text{sum}} = \frac{T_{\text{sum}}}{T_{\text{sumTotal}}}$$

with: T_{sum} , temperature sum in °C days, representative of the development stage of the crop

$$l_1 = \ln\left(\frac{fT_{\text{sum1}}}{fLAI_1} - fT_{\text{sum1}}\right) + l_2 \times fT_{\text{sum1}}$$

$$l_2 = \frac{(\ln\left[\frac{fT_{\text{sum1}}}{fLAI_1} - fT_{\text{sum1}}\right] - \ln\left[\frac{fT_{\text{sum2}}}{fLAI_2} - fT_{\text{sum2}}\right])}{fT_{\text{sum2}} - fT_{\text{sum1}}}$$

After anthesis :

$$fLAI_{\text{max}} = \frac{(1 - fT_{\text{sum}})^2}{(1 - fT_{\text{sum sen}})^2}$$

To guarantee sufficient biomass:

$$(5) LAI = \min(fLAI_{\text{max}} \times LAI_{\text{max}}, [W_{\text{total}} - W_{\text{root}}] \times S_{\text{LA}})$$

with W_{total} and W_{root} standing total and standing root biomass, respectively, in g C m⁻².

(b)

Equations for the detailed biomass production approach

$$(6) A_{gd} = \left(\frac{J_e + J_c - \sqrt{(J_e + J_c)^2 - 4 \times \theta \times J_e \times J_c}}{2 \times \theta} \right) \times d_{\text{length}}$$

with:

$$(7) J_e = \frac{C_1 \times C_q \times R_{dr} \times 0.415 \times (1 - e^{-k \times LAI})}{d_{\text{length}}}$$

$$(8) J_c = \frac{C_2 \times V_m}{24}$$

$$\text{with } V_m = \left(\frac{1}{b} \right) \times \left(\frac{C_1}{C_2} \right) \times [(2 \times \theta - 1) \times s - (2 \times \theta \times s - C_2) \times \sigma] \\ \times R_{dr} \times 0.415 \times (1 - e^{-k \times LAI}) \times C_q$$

with R_{dr} , daily incoming radiation ($\text{MJ m}^{-2} \text{d}^{-1}$), d_{length} , length of day (h), V_m Rubisco capacity ($\text{g C d}^{-1} \text{m}^{-2}$)

$$(9) C_1 = \phi_{TC3} \times C_{\text{mass}} \times \alpha_{C3} \times \left(\frac{p_i - \Gamma_*}{p_i + 2 \times \Gamma_*} \right) \text{ with } p_i = \lambda_{\text{max}} \times c_a \times P$$

$$(10a) C_2 = \frac{p_i - \Gamma_*}{p_i + K_C \times \left(1 + \frac{O_2}{K_O} \right)} \text{ with (10b)}$$

$$K_i = K_{25} \times Q_{10}^{(T-25)/10} \text{ with } i \text{ either C or O}$$

$$(11) \Gamma_* = \frac{[O_2]}{2 \times \tau}$$

with ϕ_{TC3} , a temperature stress factor (-), C_{mass} , atomic mass of carbon (g mol^{-1}), p_i , internal partial pressure of CO_2 (Pa), Γ_* , CO_2 compensation point ($\mu\text{mol mol}^{-1}$), and $[O_2]$, partial pressure of oxygen (Pa)

$$(12) A_{nd} = A_{gd} - R_d \text{ with (13) } R_d = b \times V_m$$

$$\text{and } \sigma = \left[1 - \frac{(C_2 - s)}{(C_2 - \theta \times s)} \right]^{1/2} \text{ and } s = \frac{24}{d_{\text{length}}} \times b$$

$$(14) NPP = A_{nd} - R_{\text{root}} - R_{\text{so}} - R_{\text{pool}} - R_g$$

R_i ($\text{g C d}^{-1} \text{m}^{-2}$), the maintenance respiration of i corresponding to roots, storage organs and a reserve pool, respectively, and R_g the growth respiration as

$$R_g = \max[0, 0.25 \times (A_{nd} - R_{\text{root}} - R_{\text{so}} - R_{\text{pool}})]$$

Equations for the summarized biomass production approach

$$(15) NPP = RUE \times PARINT \text{ with } PARINT = R_{dr} \times 0.5 \times (1 - e^{-k \times LAI})$$

Table 4.2. Key parameters of the approaches applied and their default values.

Symbol	Description	Default value (unit)	Source
Common parameters			
C_f	Conversion from carbon to dry matter	0.46 g C (g DM) ⁻¹	(a)
k	Light extinction factor	0.5 (–)	(b)
S_{LA}	Specific leaf area	0.048 m ² (g C) ⁻¹	(d)
$fT_{sum\ sen}$	Fraction of the total temperature sum when senescence starts (at anthesis)	0.60 (–)	(c)
T_{base}	Physiological base temperature	0 °C	(d)
LAI dynamics approaches			
<i>Parameters for the detailed LAI dynamics approach</i>			
R_g	Maximum relative growth rate of leaf area index during the juvenile stage	0.009 (°Cd) ⁻¹	(d)
$LAI_{initial}$	Initial leaf area index	0.012 m ² m ⁻²	(d)
$LAI_{juvenile\ stage}$	Threshold of leaf area index when juvenile stage ends	0.75 m ² m ⁻²	(d)
R_{d-sh}	Relative death rate due to shading	0 - 0.03 d ⁻¹	(d)
R_{d-ag}	Relative death rate due to ageing (temperature dependent)	0.03 - 0.09 d ⁻¹	(d)
$LAI_{critical}$	Critical leaf area index above which self-shading is considered	4.0 m ² m ⁻²	(d)
<i>Parameters for the summarized LAI dynamics approach</i>			
$fLAI_1$ and $fLAI_2$	Fraction of leaf area index at specific points on the leaf area development curve corresponding to specific development stages	0.05 and 0.95 (–)	(c)
fT_{sum1} and fT_{sum2}	Fraction of temperature sum at specific points on the leaf area development curve corresponding to specific development stages	0.15 and 0.50 (–)	(c)
LAI_{max}	Maximum leaf area index	5.0 m ² m ⁻²	(c)

Symbol	Description	Default value (unit)	Source
Biomass production approaches			
<i>Parameters for Farquhar photosynthesis approach (C_3 plants)</i>			
K_{25} and Q_{10}	The value of the parameter at 25 °C and the relative change in the parameter for a 10 °C change in temperature, respectively		
K_C	Michaelis constant for CO_2	30 Pa ($Q_{10}= 2.1$)	(b)
K_O	Michaelis constant for O_2	30 kPa ($Q_{10}= 1.2$)	(b)
τ	CO_2/O_2 specific ratio	2600 $\mu\text{mol } \mu\text{mol}^{-1}$ ($Q_{10}= 0.57$)	(b)
α_{C3}	C_3 quantum efficiency	0.08 $\mu\text{mol } \mu\text{mol}^{-1}$	(b)
b	R_d/V_m ratio for C_3 plants	0.015	(b)
O_2	Partial pressure of O_2	20.9 kPa	(b)
C_q	Conversion factor for solar radiation at 550 nm from $\text{MJ m}^{-2} \text{d}^{-1}$ to $\text{mol m}^{-2} \text{d}^{-1}$	4.6×10^{-3}	
θ	Co-limitation parameter	0.7 (–)	(b)
λ_{max}	Optimal ratio of intercellular to ambient CO_2 concentration	0.8 (–)	(e)
c_a	Ambient CO_2 concentration	341 $\mu\text{mol mol}^{-1}$	
P	Atmospheric pressure	100 kPa	(b)
<i>Parameters for the radiation use efficiency approach</i>			
RUE	Radiation use efficiency based on PAR and total biomass	1.38 g C MJ^{-1}	

(a) Goudriaan and Van Laar, 1994 ;(b) Haxeltine and Prentice, 1996a&b ;(c) derived from Neitsch et al., 2002;
 (d) Van Keulen and Seligman, 1987; (e) Sitch et al., 2003.

2.1.1. Leaf area index dynamics

The detailed *LAI* dynamics approach is based on temperature and leaf dry matter supply driven by the development stage of the crop (i.e. phenology). During the juvenile phase, *LAI* development is governed by temperature and its effect on cell division and extension, following an exponential growth pattern (R_g , Table 4.2). Following this exponential phase, leaf area expansion is governed by the supply of dry matter (Table 4.1, Eq. 2) and is calculated by multiplying the simulated rate of increase in leaf weight ($\frac{dW_{\text{leaf}}}{dt}$, g C m⁻² d⁻¹), based on the total amount of biomass produced multiplied by a leaf biomass allocation factor, with the specific leaf area (S_{LA} , Table 4.2). Finally, leaves senesce (Table 4.1, Eq. 3) due to (i) self-shading (R_{d-sh} , Table 4.2) when *LAI* reaches a critical value ($LAI_{critical}$, Table 4.2) and (ii) ageing after anthesis (with time of anthesis defined by $fT_{\text{sum sen}}$, Table 4.2). The relative rate at which leaves age depends on temperature (R_{d-ag} , Table 4.2), increasing with increasing temperature. This approach is used in the LINTUL model (Light, INTerception and UtILization, Spitters and Schapendonk, 1990).

The summarized *LAI* dynamics approach is governed by the development stage of the crop. Leaf area development is calculated on the basis of a forcing function, comprising a sigmoid and a quadratic component (Table 4.1, Eq. 4). *LAI* at any point in time is calculated as a fraction of an exogenously defined LAI_{max} (Table 4.2) and two shape coefficients l_1 and l_2 (-) (Eq. 4). These coefficients are defined by the fractions $fLAI_1$ and $fLAI_2$ of the maximum *LAI* (Table 4.2), and the associated fractions of the temperature sum $fT_{\text{sum}1}$ and $fT_{\text{sum}2}$ (Table 4.2), representing points on the leaf area index versus development stage curve (Neitsch et al., 2005) at specific development stages (e.g. end of juvenile stage, anthesis). The start of *LAI* senescence is defined by $fT_{\text{sum sen}}$, which is the fraction of the total growth cycle temperature sum at which senescence starts to exceed the formation of new leaf tissue. In agreement with what is applied in the more mechanistic approach, we set this starting point at anthesis. Finally, in this approach, potential *LAI* is reduced if the required biomass to support the calculated *LAI* is not available (Table 4.1, Eq. 5). This approach is applied in the LPJmL model (Lund-Potsdam-Jena managed Land, Bondeau et al., 2007) and is derived from the SWAT model (Soil and Water Assessment Tool model, Neitsch et al., 2005).

2.1.2. Biomass production

The detailed approach to describe the production of biomass is based on the description of the photosynthesis and respiration processes according to Farquhar et al. (1980) with simplifications introduced by Collatz et al. (1991; 1992). The

assimilatory process includes the conversion of CO₂ into carbohydrates. Daily *gross* photosynthesis (A_{gd} , g C m⁻² d⁻¹) is defined as a gradual transition between two limiting rates (Table 4.1, Eq. 6). Photosynthesis is determined either by the amount of intercepted light (J_e , g C m⁻² h⁻¹, Table 4.1, Eq. 7) or by the available amount of the enzyme Rubisco (J_c , g C m⁻² h⁻¹, Table 4.1, Eq. 8). Those rates are both influenced by ambient temperature (Table 4.1, Eqs. 9, 10a/b and 11), via τ (Table 4.2) and via K_C and K_O , the temperature-dependent Michaelis-Menten constants for CO₂ and O₂ (Table 4.2), respectively. Daily *net* photosynthesis (A_{nd} , g C m⁻² d⁻¹, Table 4.1, Eq. 12) is calculated as daily gross photosynthesis minus “dark” respiration (R_d , g C m⁻² d⁻¹). R_d is scaled to the maximum catalytic capacity of Rubisco per unit leaf area (V_m , g C m⁻² d⁻¹, Table 4.1, Eq. 13). To calculate net primary production (NPP , g C m⁻² d⁻¹, Table 4.1, Eq. 14), maintenance respiration for the various organs (R_i , g C m⁻² d⁻¹, for roots, storage organs, and carbohydrate pool, respectively) is subtracted from daily net photosynthesis and 25% of the remaining assimilates is assumed to be expended in growth respiration. This approach is used in various models, but the present equations (Haxeltine and Prentice, 1996) are implemented within the LPJmL model (Bondeau et al., 2007).

Alternatively, the summarized approach is based on a linear relationship between accumulated intercepted radiation, and accumulated biomass over the growing season. The slope of this linear relation is called radiation use efficiency (RUE , Table 4.2) (Monteith, 1977) and summarizes the combined effect of photosynthesis and respiration processes. The product of the daily intercepted amount of photosynthetic active radiation (PAR_{INT}) and RUE gives the net increase in biomass over the day (Table 4.1, Eq. 15). This approach is used in models such as LINTUL (Spitters and Schapendonk, 1990), CropSyst (Stöckle et al., 2003) and CERES (Ritchie and Otter, 1985).

These four approaches (two for LAI dynamics and two for biomass production) were combined in various ways resulting in four models (Table 4.3).

Table 4.3. Overview of combination of processes and their derived models.

Light utilization approach: Biomass production	Light interception approach: <i>LAI</i> dynamics	Model name
Farquhar photosynthesis combined with:	detailed	Detailed crop model
	summarized	LPJmL (Lund-Potsdam- Jena managed Land)
<i>RUE</i> combined with:	detailed	LINTUL (Light, INTERception and UtiLization)
	summarized	Summarized crop model

2.2. Model testing

To test the four models, measured *LAI* and biomass data for spring wheat, under optimal agronomic conditions for potential growth, from contrasting locations, were collected with their associated weather data: Australia (Meinke et al., 1997), Europe (Van Oijen et al., 1998; Bender et al., 1999; Ewert and Pleijel, 1999; Van Oijen and Ewert, 1999), and USA (Kimball et al., 1995; 1999; Ewert et al., 2002). Details are given in Table 4.4. The locations vary in temperature conditions during the growing season: in the USA temperatures (i.e. number of days > 22.5 °C) are higher during the end of the growing season than in the Netherlands or Australia. Moreover, radiation levels during the growing season are higher in the USA than in the other locations.

The four models were calibrated with respect to phenology, *LAI* dynamics and yield for these locations. The parameters $fT_{\text{sum sen}}$, S_{LA} , LAI_{max} and total temperature sum (T_{sum}) of the growth cycle were first estimated from the observed data, and subsequently adjusted according to model results (i.e. simulated *LAI* and yields). The calibration was done manually by trial-and-error method, comparing the model results with the observed results. R_g was calibrated on the basis of model results, guided by values found in the literature. For the biomass production approaches, *RUE* was directly estimated from the observed data: it was not calibrated, to avoid the compensation effect with the calibration on S_{LA} .

To evaluate the quality of the model outputs, we used the relative root mean square error ($rRMSE^1$, Wallach et al., 2006) for yield and the relative mean absolute error ($RMAE^2$, Mayer and Butler, 1993) for *LAI* dynamics.

$$^1 rRMSE = \sqrt{\frac{\sum_{i=1}^n (S_i - O_i)^2}{n}} \times \frac{1}{O_i} \text{ with } i \text{ (number of locations)} = 1, \dots, n$$

$$^2 RMAE = \frac{1}{n} \sum_{i=1}^n \left(\frac{|S_i - O_i|}{S_i} \right) \text{ with } i \text{ (number of } LAI \text{ observations over the growing cycle)} = 1, \dots, n$$

Table 4.4. Weather characteristics for the various locations of experimental datasets.

	Australia, Queensland	Europe, The Netherlands	USA, Arizona			
	1993	1995	1996	1992-93	1993-94	1994-95
Average temperature during the growing season (°C)	15.5	14.8	13.4	15.2	14.7	15.5
Days < 7.5°C	2	7	12	1	9	7
Days > 22.5°C	4	9	3	10	10	16
Total radiation (MJ m ⁻² growing season ⁻¹)	2033	2042	1998	2579	2904	2649
Intercepted PAR (MJ m ⁻² growing season ⁻¹)	554	669	616	823	797	724

2.3. Systematic comparison of model behaviour to climate variability

To investigate the relative importance of the two growth processes on simulated crop yield and their ability to capture climatic variability, the models were run with weather data representing a wide range of climatic conditions in Europe (Figure 4.1). Assessing model behaviour for a wide range of environmental conditions should demonstrate how robust the different approaches are under different conditions and therefore how suitable the different approaches are for application at larger scales. Nine locations were selected across Europe: Denmark, the United Kingdom, the Netherlands, Germany, France (centre and south), Spain (centre and south), and Italy. They represent the European climatic gradient according to the classification from Metzger et al. (2005). Daily data for minimum and maximum temperature and incoming short-wave radiation for the year 1982 (for this specific year daily weather data were available for the nine locations) were extracted from a database described by Van Kraalingen et al. (1991). In addition to location-specific weather data, the models were run with default parameters for *LAI* dynamics and biomass production approaches as obtained from the calibration step for the Netherlands (assuming those parameters to be representative for Europe). We adapted the phenology parameters for each location. As sowing and harvesting dates for spring wheat were not available for all locations, we used data for spring barley as a proxy (Table 4.5, Boons - Prins et al., 1993).



Figure 4.1. Location of the nine weather stations, representing a climatic gradient (from Denmark to southern Spain).

To evaluate the sensitivity of simulated yields to the modelling approaches, we performed an analysis of variance (ANOVA) to identify significant differences among simulated yields associated with the different approaches and locations (assumptions of ANOVA of normality of the data and homogeneity of variance were not violated). With respect to location, we classified the 9 locations in two categories (North vs. South, using latitude as a criterion, Table 4.5) to be able to run the ANOVA and identified whether simulated yields significantly differed among the locations, but also between the different approaches (interaction effect) in the different locations.

The outcomes of this analysis indicate, as a first step, whether the simulated yields differ significantly among approaches and locations. But, they do not identify which process most strongly affects the simulated yield variability across locations. Therefore, we used the relative standard deviations (RSD^3) (i) to determine if location-specific weather influenced the outcomes of a certain approach and thus if there is any effect of climatic variability on model outcomes and (ii) to understand the relative importance of the different light interception (i.e. LAI dynamics) and utilization (i.e. biomass production) approaches to capture this climatic variability. RSD was calculated for (i) RUE , representing the light utilization approach, (ii) intercepted photosynthetically active radiation (PAR), representing the light interception approach and (iii) yield, which is the variable of interest and the integrated result of both processes.

³ $RSD = \frac{\sigma_{yield}}{yield}$

Finally, a sensitivity analysis was carried out on parameters of the summarized and detailed approaches of LAI dynamics to evaluate the impact of a change in a given parameter on simulated yields. Parameter values given by the initial approach (i.e. Table 4.2) were used as default. We applied a range of variation in the parameters based on the variation found in the observed data and in the literature. LAI reached up to $7 \text{ m}^2 \text{ m}^{-2}$ in the observed data. Furthermore, Hay and Porter (2006) indicate that 90% of the incoming radiation is usually intercepted at a LAI varying from 3 to $5 \text{ m}^2 \text{ m}^{-2}$ (Hay and Porter, 2006). Therefore, we tested the sensitivity for this parameter from 3 to $7 \text{ m}^2 \text{ m}^{-2}$. S_{LA} varies from 0.036 to $0.061 \text{ m}^2 (\text{g C})^{-1}$ in the observed data, in agreement with Stöckle et al. (2003). The first phase in LAI development, critical for LAI dynamics, is defined by the parameter R_g in the detailed approach. From calibration and values given in Van Delden et al. (2001) for spring wheat, we defined a range from 0.005 to $0.013 (\text{°C d})^{-1}$. Finally, the timing of the onset of leaf senescence is defined by the parameter $fT_{\text{sum sen}}$ which varied from 0.5 to 0.7 in the observed data

We varied each parameter within the defined range (Table 4.6) by small increments of $\pm 1\text{-}4\%$, depending on parameter. We considered small increments to be able to identify the sensitivity to each parameter. The sensitivity index S_i ⁴ is based on the local variation in the output value with respect to the variation in a given parameter (Wallach et al., 2006). If S_i is small ($S_i < 0.5$), the simulated yield is not very sensitive to the parameter tested. This analysis gives some indication of the relative importance of the parameter for different locations and different approaches.

Table 4.5. Location-specific phenological cultivar parameters used for the systematic comparison of models.

Countries	Symbol	Latitude	Location*	Emergence (day of year)	Temperature sum till maturity (°Cd)
Denmark	DK	57.1	north	90	1577
United Kingdom	UK	52.35	north	51	1693
The Netherlands	NL	52.1	north	85	1924
Germany	GE	48.12	north	60	1383
France (centre)	Fr_centre	47.97	north	69	1657
France (south)	Fr_south	43.62	south	36	2149
Italy	IT	42.42	south	31	2044
Spain (centre)	SP_centre	40.45	south	31	2022
<i>Spain (south)</i>	<i>SP_south</i>	<i>37.42</i>	<i>south</i>	<i>31</i>	<i>2443</i>

* regions at latitudes below 45° are considered to be southern.

$$^4 S_i = \frac{\frac{\delta_{\text{yield}}}{\text{yield}}}{\frac{\delta_{\text{parameter}}}{\text{parameter}}}$$

Table 4.6. Statistical setting used in the sensitivity analysis on the key parameter values of the light interception approaches.

Symbol	Description	Unit	Default	Standard deviation*	Increment	Source*
S_{LA}	Specific leaf area	$m^2 gC^{-1}$	0.048	0.0125	0.001	Dataset and Stöckle et al. (2003)
R_g	Relative growth rate of leaf area index during the juvenile phase	$^{\circ}C^{-1}d^{-1}$	0.009	0.004	0.0004	Van Delden et al. (2001)
LAI_{max}	Maximum leaf area index	$m^2 m^{-2}$	5	2	0.1	Dataset and Hay and Porter (2006)
$fT_{sum\ sen}$	Fraction of the total temperature sum when senescence starts	Unitless	0.6	0.1	0.01	Dataset and Neitsch et al. (2002)

* derived from the observed data and literature.

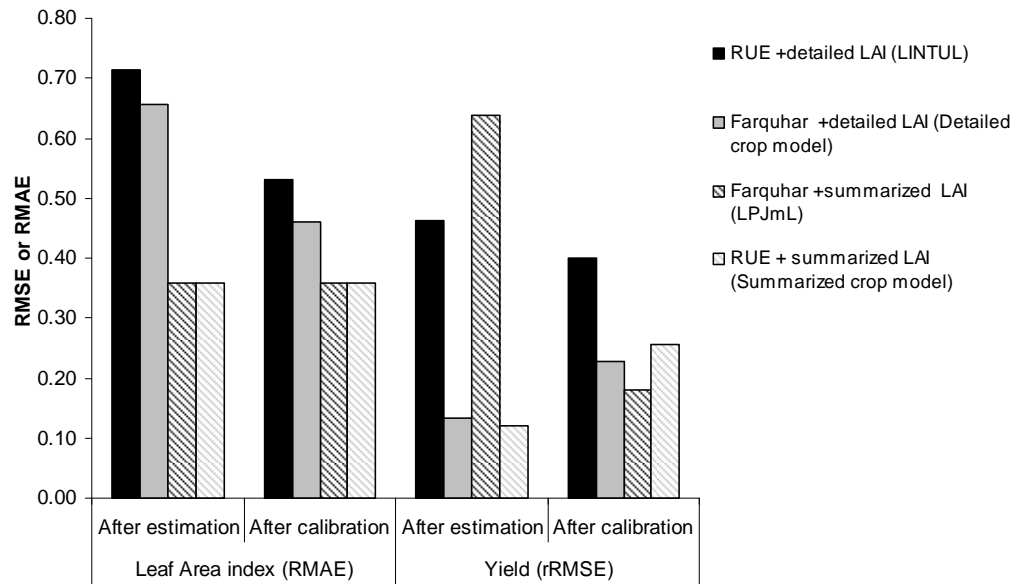


Figure 4.2. Relative mean absolute error (RMAE) to analyze the performance of the LAI dynamics simulation and relative root mean square error (rRMSE) to analyze the performance of simulated yield for the four studied models.

3. Results

3.1. Test of the models with experimental data

After calibration, simulated yields reproduced the observed yield with a $rRMSE$ ranging between 18% and 40% (Figure 4.2), depending on the model. Agreement between simulated and observed yields is closest for models using the summarized LAI dynamics approach, especially when parameter values estimated from the observed data were used. The simulations of LAI dynamics support this observation. The model including the detailed LAI dynamics approach performs least satisfactorily with a $RMAE$ of 0.50 vs. a $RMAE$ of 0.36 for the models with the summarized LAI dynamics approach.

On the one hand, it is clear that by estimating LAI_{max} from the observed data, LAI is simulated satisfactorily in the summarized LAI dynamics approach ($RMAE = 0.36$), with no improvement through calibration (Figure 4.2). On the other hand, calibration is important in the detailed LAI dynamics approach, especially for R_g . The default value of R_g (0.009 ($^{\circ}Cd$) $^{-1}$) is based on Van Keulen and Seligman (1987) who included a wide range of data from different wheat cultivars (i.e. both spring and winter wheat) in their analysis. Spring wheat requires a higher value of R_g , corresponding to the range observed by Van Delden et al. (2001).

Finally, calibration of S_{LA} and $fT_{sum\ sen}$ also improves the simulated yields, independent of the LAI dynamics approach (Eqs. 2 and 5). With respect to the biomass production approaches, a lower value of RUE was estimated from the data for locations with higher temperatures and total accumulated radiation over the growing cycle (i.e. USA, Table 4.7). However, because of lack of data, it was not possible to define a significant relationship for RUE as a function of radiation and temperature from our dataset, as the data were too limited. Table 4.7 gives the calibrated parameters for each location.

Table 4.7. Parameter values after calibration using experimental datasets.

	Australia (Queensland)	Europe (the Netherlands)	USA (Arizona)
$fT_{sum\ sen}$	0.61	0.54	0.55
P_{hu}	1804	1609	2070
R_g	0.013	0.013	0.010
S_{LA}	0.06	0.045	0.054
LAI_{max}	7	6.75	6.35
RUE^*	1.52	1.33	1.01

* RUE was not calibrated to avoid a compensation error with the S_{LA} parameter.

The models with the summarized *LAI* dynamics approach could simulate crop productivity reasonably well in locations with different radiation intensities and temperature conditions (e.g. USA, Arizona vs. Europe, the Netherlands). However, this is achieved only when applying parameter values estimated from the dataset. When applying the models for a wider range of conditions, the issue of data availability needs to be considered.

3.2. Model behaviour in response to climatic conditions

3.2.1. Sensitivity of simulated yield to the different modelling approaches

To investigate the behaviour of the models in capturing the effects of different climatic conditions as relevant for regional applications, we ran the four models with weather data from a climatic range across Europe. The parameter values (except for phenology) were calibrated for the Netherlands, considered representative for Europe from our dataset. Simulated yields vary from a maximum of 8.38 Mg dry matter ha⁻¹ in southern Spain to a minimum of 4.44 Mg dry matter ha⁻¹ in Germany (Figure 4.3). The minimum yield was simulated with the combination of the detailed *LAI* dynamics and the *RUE* approach (LINTUL), while the maximum was simulated with the combination of the summarized *LAI* dynamics and the *RUE* approach (summarized crop model). LINTUL shows the strongest response to climatic variability ($RSD = 0.20$, Table 4.8), while the two models using the Farquhar approach (LPJmL and detailed crop model) show the weakest response ($RSD = 0.12$, Table 4.8).

Further, to better understand which process is more sensitive to climatic variability, we used the relative standard deviations (RSD) of *RUE* and intercepted *PAR* (Table 4.8). The calculated *RUE* value, based on outcomes from the Farquhar photosynthesis approach, is slightly influenced by climatic variability ($RSD = 0.04$). Intercepted *PAR* shows the highest RSD values, especially when using the detailed *LAI* dynamics approach, demonstrating that this process is most sensitive to climatic variability ($RSD = 0.24 - 0.26$).

Table 4.8. Relative standard deviation (*RSD*) to define the ability of each approach to capture climatic variability.

	Summarized		Detailed	
	<i>LAI</i> dynamics approach		<i>LAI</i> dynamics approach	
	Summarized		Detailed crop	
	LPJmL (Farquhar)	crop model (<i>RUE</i>)	model (Farquhar)	LINTUL (<i>RUE</i>)
<i>RSD</i> yield	0.12	0.15	0.12	0.20
<i>RSD</i> intercepted PAR	0.16	0.16	0.24	0.26
<i>RSD RUE</i>	0.04	0.00	0.04	0.00

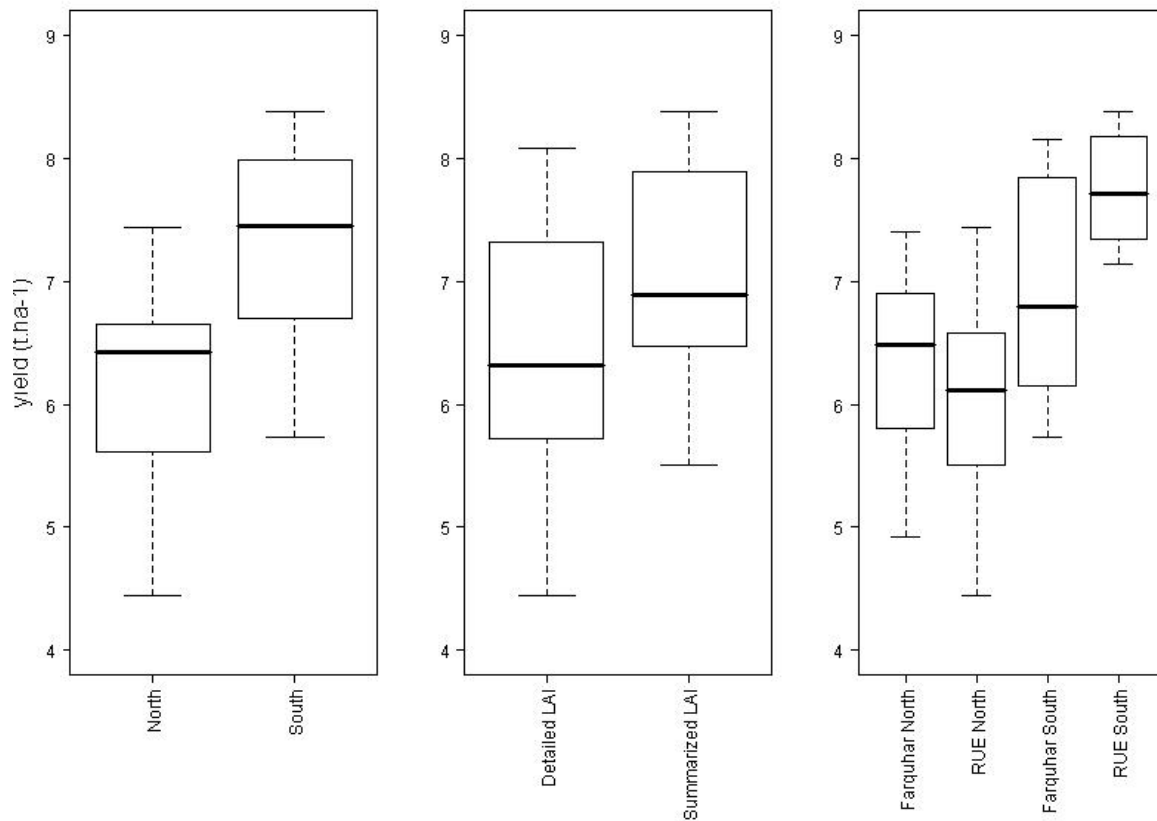


Figure 4.3. Range of simulated yields for a wide range of conditions in Europe according to:
 (a) location (i.e. North vs. South),
 (b) *LAI* dynamics approach per location (i.e. detailed *LAI* vs. summarized *LAI*), and
 (c) biomass production approach depending on locations (i.e. Farquhar North vs. *RUE* North and Farquhar South vs. *RUE* South).

Table 4.9. Analysis of variance to identify whether the different modelling approaches and locations (North vs. South) result in significant differences in simulated yields.

Response: YIELD	Sum of squares	df	Fvalue	Pr(>F)	
Biomass production approach	0.30	1	0.56	0.46	
LAI dynamics approach	4.23	1	7.99	0.009	**
Location (North vs. South)	12.78	1	24.15	3.50×10^{-5}	***
Biomass production \times LAI dynamics approach	0.47	1	0.89	0.35	
Biomass production approach \times location	2.93	1	5.53	0.03	*
LAI dynamics approach \times location	1.44	1	2.72	0.11	
Biomass production approach \times LAI dynamics approach \times location	0.91	1	1.73	0.20	
Residuals	14.82	28			

Significance codes: '***' 0.001; '**' 0.01; '*' 0.05

From the ANOVA (Table 4.9), we identify that location (i.e. northern vs. southern regions) has a clear influence on simulated yields, independent of the approach chosen ($p < 0.001$): simulated yields are higher in southern regions than in northern regions (Figure 4.3a). Moreover, the choice of the LAI dynamics approach is important which is independent of the location ($p < 0.01$): simulated yields are always higher with the summarized LAI dynamics approach (Figure 4.3.b).

Finally, the ANOVA demonstrates a significant difference in simulated yields for the two biomass production approaches, depending on location ($p < 0.05$): RUE simulates higher yields in southern regions than the Farquhar approach, while the Farquhar approach simulates higher yields in northern regions (Figure 4.3c). The latter result supports our observations that if the RUE approach is used, the value of RUE should be adjusted to temperature and radiation conditions.

3.2.2. Sensitivity to parameter values for the LAI dynamics approaches

The choice of the light interception approach has a significant influence on simulated yields (ANOVA results, Table 4.9) and this process most strongly reflects the effect of climatic variability (higher RSD) on yields. Subsequently, we carried out a sensitivity analysis on key parameters of the light interception approaches to assess their relative importance for the simulated yields, when combined with the RUE approach. Figure 4.4 shows the sensitivity index (S_i) for the parameters tested in the detailed and

summarized *LAI* dynamics approaches. The sensitivity of simulated yield is different for the different parameters considered (Figure 4.4). The sensitivity of simulated yields is irregular in the detailed *LAI* dynamics approach, while it is smooth in the summarized *LAI* dynamics approach.

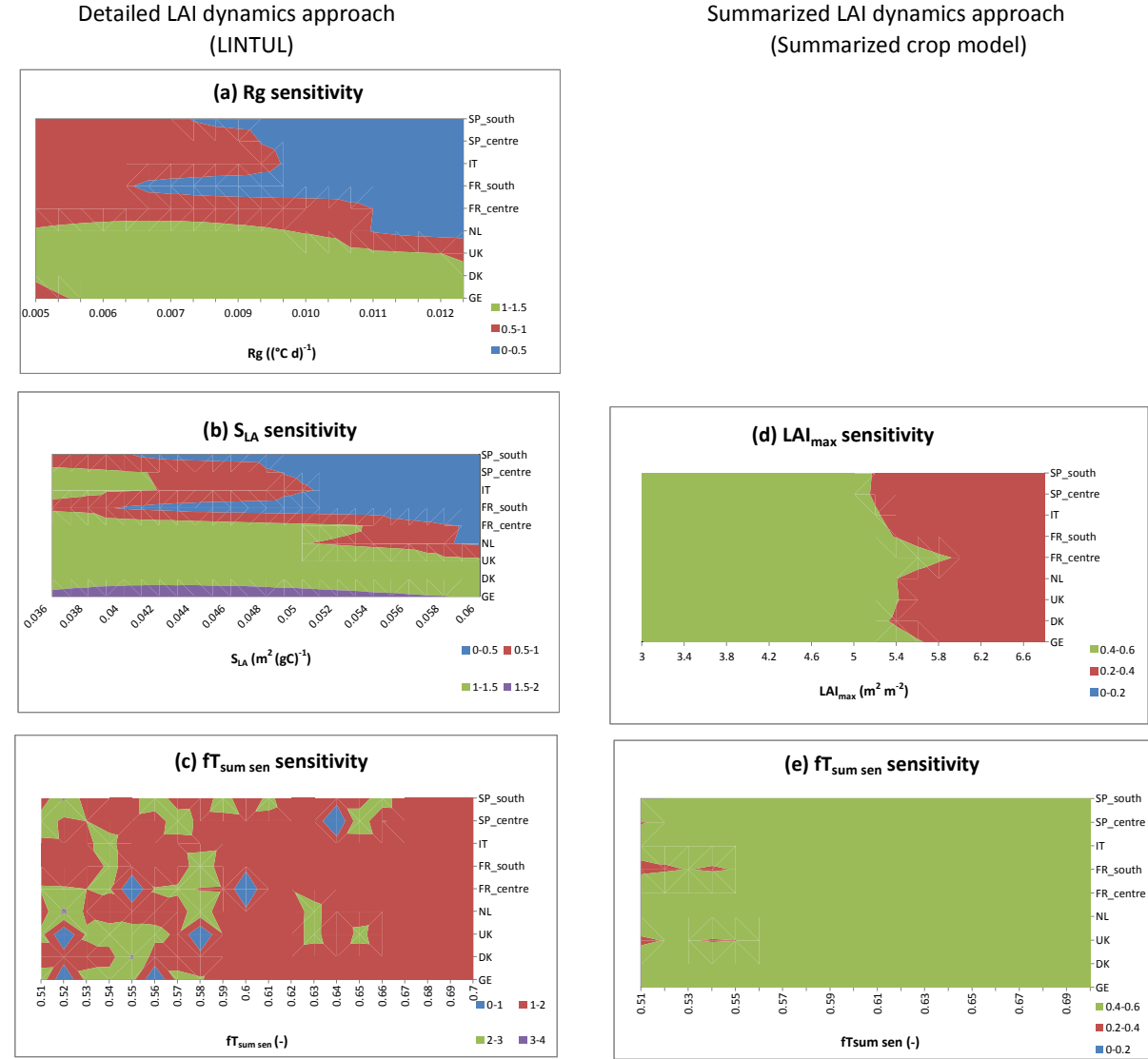


Figure 4.4. Sensitivity index (S_i) of the yield for the main parameters according to the *LAI* dynamics approaches. $S_i = 1$ means that a change in the parameter value will induce the same relative change in simulated yield. Symbols referring to the locations are the same as those in Table 4.5. The parameters tested are:

- (a) R_g : Relative growth rate of leaf area index during the juvenile phase
- (b) S_{LA} : Specific leaf area
- (c and e) $ft_{sum\ sen}$: Fraction of the total temperature sum when senescence starts
- (d) LAI_{max} : Maximum leaf area index

LINTUL (Light, *INT*erception and *UTi*lization) is the model combining the *RUE* approach with the detailed *LAI* dynamics approach, while summarized crop model is the one combining the *RUE* approach with the summarized *LAI* dynamics approach.

Note that the colours in the various figures represent different ranges of S_i

The relative growth rate of LAI in the exponential phase (R_g), used in the detailed LAI dynamics approach, was varied from 0.005 to 0.013 ($^{\circ}\text{C d}^{-1}$). The sensitivity of simulated yield to this parameter is variable, depending on location and the value of the parameter itself. On the one hand, simulated yields are highly sensitive to R_g in the northern locations (Germany, Denmark, the United Kingdom and partly the Netherlands) with a S_i between 1 and 1.5 (i.e. a change in R_g by 1% will result in a change in simulated yield of 1% to 1.5%). On the other hand, in the southern regions, a change in R_g has a relatively smaller impact on simulated yields, especially for values exceeding the default value of 0.009 ($^{\circ}\text{C d}^{-1}$) (Figure 4.4a). From the calibration we obtained a value for R_g of 0.013 ($^{\circ}\text{C d}^{-1}$), which implies that only simulated yields in Germany, Denmark and the United Kingdom will be highly sensitive to a variation in this value ($S_i > 1$).

S_{LA} (varying from 0.0036 to 0.06 $\text{m}^2 (\text{g C})^{-1}$) behaves similarly, with the exception of a higher sensitivity ($S_i > 1$) in southern regions (i.e. Italy and Central Spain) with values of $S_{LA} < 0.042 \text{ m}^2 (\text{g C})^{-1}$ (Figure 4.4b). When using a value of 0.045 $\text{m}^2 (\text{g C})^{-1}$ for S_{LA} (as derived from the calibration for the Netherlands), the northern regions will be highly sensitive ($S_i > 1$) and most of the southern regions moderately sensitive ($0.5 < S_i < 1$), except for the South of France and South of Spain, where yield sensitivity is relatively small ($S_i < 0.5$) to variation in this parameter.

Finally, the sensitivity of simulated yield to $fT_{\text{sum sen}}$ in the detailed LAI dynamics approach is high (Figure 4.4c). For values of $fT_{\text{sum sen}}$ exceeding 0.66, the sensitivity of simulated yield is uniform among locations, i.e. independent of changes in $fT_{\text{sum sen}}$, simulated yield will at least change by the same proportion ($S_i > 1$). For values of $fT_{\text{sum sen}}$ below 0.6, the yield is highly sensitive in many locations ($S_i > 2$).

Interestingly, for a few locations and some specific values of $fT_{\text{sum sen}}$, the simulated yield is not sensitive to a change in its value (e.g. central Spain for a value of 0.64 and the United Kingdom for a value of 0.52). For a $fT_{\text{sum sen}}$ value of 0.54 (the calibrated value for the Netherlands), simulated yield is highly sensitive to variations in its value, with S_i varying from 1.55 to 2.45 for any location. For the model using the summarized LAI dynamics approach, the results of the sensitivity analysis are much more straightforward, with a moderate sensitivity ($0.4 < S_i < 0.6$) of the simulated yield to both parameters LAI_{max} (Figure 4.4d) and $fT_{\text{sum sen}}$ (Figure 4.4e). For LAI_{max} exceeding 5.5 $\text{m}^2 \text{m}^{-2}$, the sensitivity of the simulated yield is even lower ($0.2 < S_i < 0.4$), independent of the location.

4. Discussion

4.1 General behaviour of the models

All models simulated higher yields in southern regions than in northern regions (Figure 4.3a), associated with longer growing seasons and higher radiation intensities during the growing season, due to earlier sowing dates and higher temperature sums till physiological maturity (Boons-Prins et al., 1993, Table 4.5). The consequence is higher accumulated intercepted *PAR*, leading to higher biomass accumulation and therefore higher simulated yields. Such model outcomes are contrary to what is observed (yields are usually lower in southern regions than in northern regions in Europe, Van Oijen and Ewert, 1999). Indeed, the input data in terms of phenology were derived from spring barley, because of lack of available data for spring wheat (Boons-Prins et al., 1993). This proxy may be questionable, as other studies report that spring wheat can be sown between November-December in Mediterranean regions (Russell and Wilson, 1994). Hence, this result underlines the importance of adequately including farmers' practices at different locations, as a response to the spatial variability in climate (Reidsma et al., 2010). Further, it may not be sufficient to only adapt model parameters for phenology (e.g. Théron et al., 2010) for larger scale applications, but also growth processes are particularly important to capture the effects of climatic variability.

4.2. Biomass production approaches

The two biomass production approaches result in significant differences in simulated yields, differentiated among locations. The *RUE* approach simulates higher yields in southern regions than the Farquhar approach, while the Farquhar approach simulates higher yields in northern regions (Figure 4.3c). Using the *RUE* approach, with a constant *RUE* value for large scale applications (Tan and Shibasaki, 2003; Liu et al., 2007), we might over-simplify, ignoring effects of high temperatures and high radiation intensities on net photosynthesis, both considered in the Farquhar approach (Figure 4.5) or on daily light use efficiency (Choudhury, 2000; 2001).

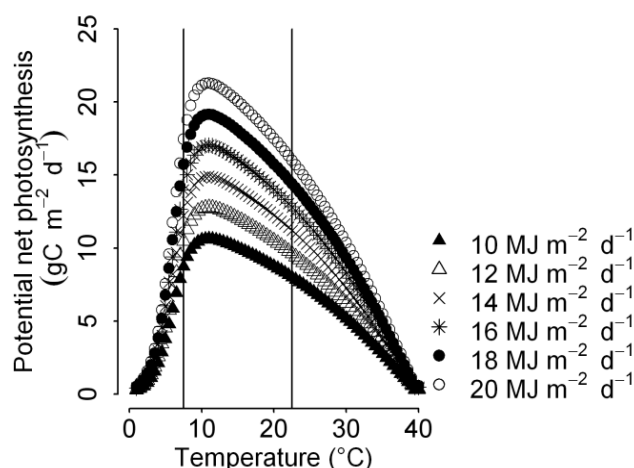


Figure 4.5. Potential net photosynthesis (A_{nd}) as a function of temperature and a number of constant total radiation intensities at one CO_2 concentration of 341 ppm.

The two vertical lines indicate the temperature limits where A_{nd} is 75% of its maximum value (7.5 and 22.5 °C) for the Farquhar approach (detailed biomass production approach).

The use of a constant RUE (over the growing cycle as a whole) incorporates a lower conversion efficiency during the grain filling period due to, mostly, reallocation of assimilates to the grains (Van Keulen and Seligman, 1987). Moreover, when LAI reaches values of 3-4 $m^2 m^{-2}$, the effect of radiation intensity on RUE reflects the light saturation effect. Higher radiation then leads to lower RUE , as the leaves at the top of the canopy are light saturated and thus, higher light absorption does not lead to higher assimilation (Hay and Porter, 2006). As a consequence, conversion efficiency calculated on the basis of total absorbed radiation is lower. Finally, this aggregated value of RUE also includes the temperature effect on respiration processes. So, to keep the model as simple as possible, i.e. to appropriately balance between data availability and model structure for large scale application (Addiscott, 1998; Hansen and Jones, 2000; Jagtap and Jones, 2002), the RUE approach could be extended by incorporating the effects of temperature and radiation (Stöckle and Kemanian, 2009). We propose to adapt the seasonal value (Medlyn, 1998; Sinclair and Muchow, 1999), using results of the Farquhar photosynthesis algorithms (Mitchell et al., 2000). Figure 4.6 demonstrates how the effect of temperature and radiation on the value of RUE could be expressed on a daily basis (derived from Choudhury, 2001) and on a seasonal basis (derived from the present study), which is more appropriate for regional applications of crop growth models.

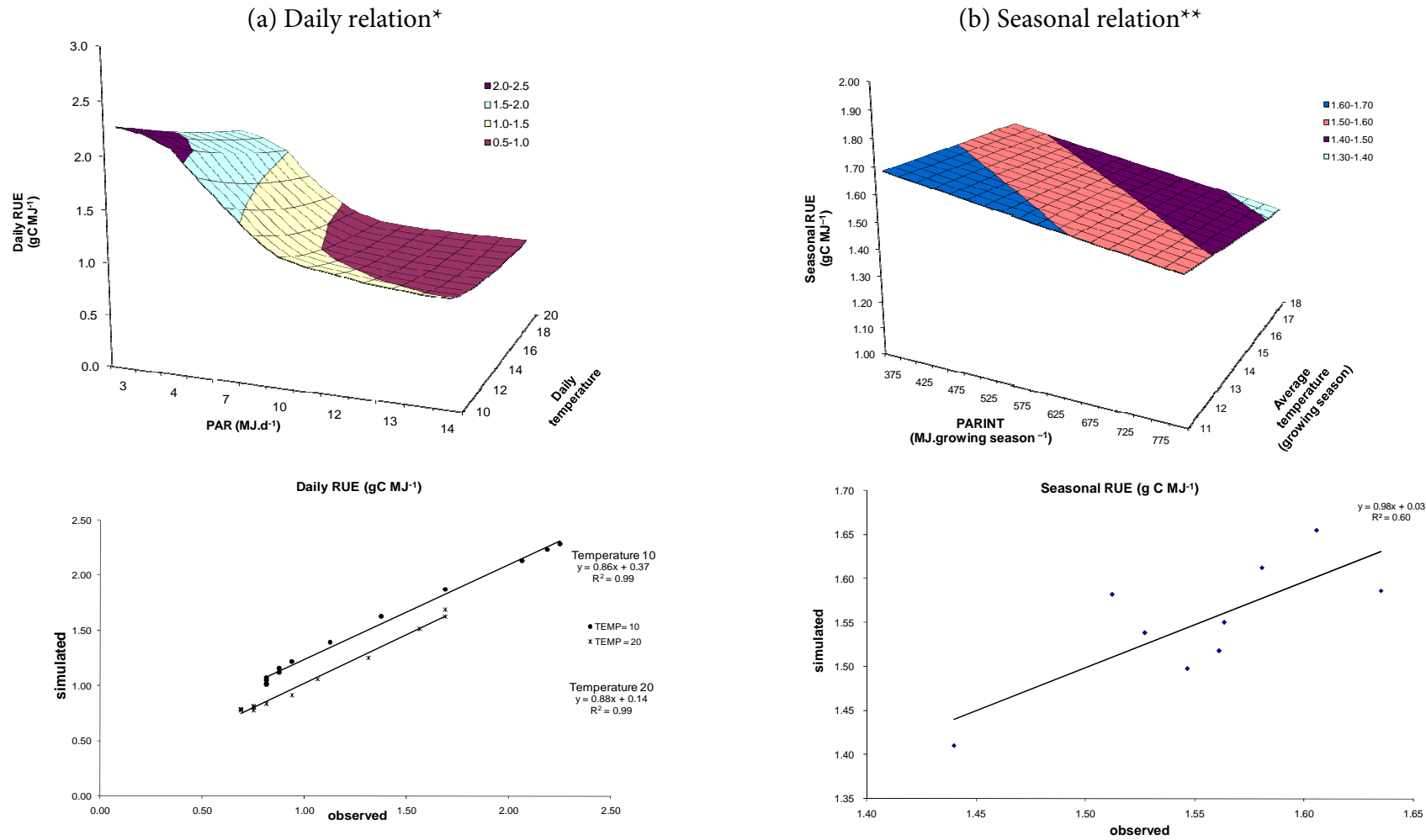


Figure 4.6. Scaling of RUE for large scale applications, from (a) an exponential relation on a daily basis (from Choudhury, 2000) (b) a linear relation on a seasonal basis: graphical representation of the relation and observed vs. simulated RUE for both cases.

* The daily RUE ranges 0.75 to 2.25 gC MJ⁻¹, according to $RUE = 0.75 + 2.5 \times \exp^{-(0.016 \times temp) \times PAR}$ with $10 \leq temp \leq 20$ and $3 \leq PAR \leq 14$

$temp$: daily average temperature (°C) and PAR : daily photosynthetically active radiation (MJ m⁻² d⁻¹).

** The seasonal ranges 1.45 to 1.65 g C MJ⁻¹ according to $RUE = 2.1 - 3.5 \cdot 10^{-4} \times PARINT - 2.5 \cdot 10^{-2} \times temp$ with $11 \leq temp \leq 18$ and $375 \leq PARINT \leq 800$

$temp$: average temperature during the growing season (°C) $PARINT$: intercepted photosynthetically active radiation (MJ m⁻² growing season⁻¹).

4.3 *LAI* dynamics approaches

The two different light interception approaches result in significant differences in simulated yields, independent of the location considered, and most clearly captures climatic variability. These results confirm earlier work that identified light interception as an important factor in determining crop growth (Heath and Gregory, 1938; Watson, 1947) and with later ones in the context of climate change studies (Ewert, 2004b; Jamieson et al., 1998).

Using the same parameter values, the summarized *LAI* dynamics approach simulates higher yields than the detailed *LAI* dynamics approach (Figure 4.3b). We assumed the $fT_{\text{sum sen}}$ parameter to have the same value in both approaches, as it was difficult to find an unambiguous definition of this parameter (i.e. onset of leaf senescence). In some cases, $fT_{\text{sum sen}}$ is equivalent with the physiological meaning of leaf senescence, i.e. when leaves actually start to senesce (Havelka et al., 1984), while in some other cases, it is a visual interpretation of the phenomenon (Mi et al., 2000; Araus and Tapia, 1987), when the death rate of leaves overrides their growth rates. We considered the timing of the onset of senescence to be equal in the two approaches, i.e. at anthesis, in line with the detailed approach. However, the original description of the summarized approach defined the timing of the onset of senescence more on a visual observation: “*LAI* will remain constant until leaf-senescence begins to exceed leaf growth” (Neitsch et al., 2005, p.294). Furthermore, we assumed the timing of the onset of senescence to be identical across locations. However, phenological characteristics (e.g. temperature sum requirements till anthesis) of wheat vary among cultivars (Slafer and Rawson, 1994), suggesting the need to also define site-specific values for $fT_{\text{sum sen}}$. Hence, looking at differences in simulated yield due to the different modelling approach, we analyzed the different responses of the models due to the use of the same parameter values and could clarify the different underlying assumptions, lumped in the $fT_{\text{sum sen}}$ value (e.g.), essential for simulated yields.

5. Conclusions

From the two main processes determining growth, i.e. light interception and utilization, we found that the first is most important in explaining yield sensitivity to climatic variability. We also showed that a different light interception approach results in significant differences in simulated yields, irrespective of the location. We conclude that for large scale application of crop models, particular attention should be given to the simulation of *LAI* and light interception. Most critical in this respect is the representation of leaf senescence, particularly the onset of senescence which is

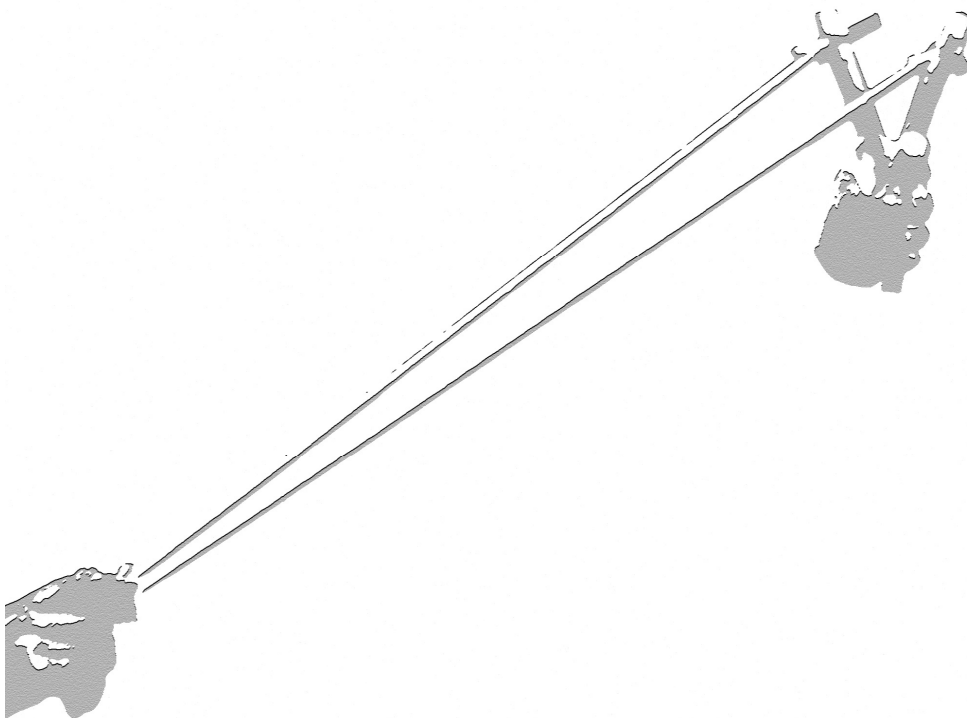
modelled differently in crop models, but has considerable impact on the simulation results.

We also found that oversimplification of processes can lead to omission of important relationships, as evident in the application of the *RUE* concept. We propose that models using the concept of *RUE* should adjust seasonal *RUE* depending on temperature and radiation.

We finally demonstrated that through an integrated use of complex (i.e. detailed) and simple (i.e. summarized) approaches more insight can be gained about how to model crop growth for large scale applications, in support of decisions on the right trade-off between data availability and model detail.

CHAPTER 5

Criteria of selection:
crop type*



* This chapter is to be submitted as :

Adam, M., Wery, J., Leffelaar, P.A., Ewert, F., Corbeels, M., Van Keulen, H. Simulating pea growth from wheat growth: a systematic approach for model re-assembly. *European Journal of Agronomy*.

Parts of this chapter were presented as :

Roux, S., Adam, M., Wery, J., 2010. An interdisciplinary framework for the development and analysis of crop models. Oral presentation at International Congress on Environmental Modelling and Software Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada. 5 - 8 July 2010.

Adam, M. , Leffelaar, P.A. , Ewert, F., Corbeels, M., Van Keulen, H., Wery, J., 2010. Developing new crop models within a flexible crop modelling framework: use of crop physiological knowledge to change a wheat model into a pea model. Poster Presentation at XIth European Society of Agronomy congress, Montpellier, France. 29 August – 03 September, 2010.

Abstract

The process of crop modelling to develop an operational software tool requires different skills, from conceptualization of the biophysical system to computer programming, involving three main scientific disciplines: agronomy, mathematics and software engineering. Model building implies transforming a conceptual model into sets of mathematical equations and then translating these equations into a computer program. The focus of this paper is on a procedure to re-assemble models, i.e. develop a new crop model from an existing one, using an existing crop modelling framework and crop physiological knowledge. Modifications to the initial crop model were classified according to three categories: (i) changes in the values of parameters, (ii) changes in equations, and (iii) changes in the overall structure of the model. We illustrate the approach with a case study transforming a wheat crop model into a pea crop model. We conclude that the use of the approach to re-assemble a new model in combination with a crop modelling framework leads to (i) integration of different disciplines around a modelling objective, (ii) combining new (expert) knowledge with existing models without ‘re-inventing the wheel’, and (iii) efficient communication with the user of the tool.

Keywords: crop modelling framework, crop physiology, agronomic model, mathematical model, software.

1. Introduction

Crop growth models can be applied in support of formalization of agronomic expert knowledge (Bouman et al., 1996) and/or of design alternative cropping systems (Bergez et al., 2010) and their assessment with respect to multiple criteria that are related to crop productivity and environmental externalities. Although crop models are often viewed as simulation software tools, they can assume different forms, from graphical representations of the cropping system (i.e. a conceptual model) to mathematical algorithms and coded computer programs. Accordingly, the process of modelling to arrive at an operational simulation tool requires combination of different skills, ranging from conceptualization of the biophysical system (system analysis and agronomy) to the coding of the mathematical equations (mathematics and software engineering). In applying crop modelling as a research tool, the researcher has three main options while starting a simulation study (Sinclair and Seligman, 1996):

- (i) *create his/her own model*: that requires a high investment, including knowledge of crop physiology and computer programming (Rabbinge and De Wit, 1989), but it has the advantages that the researcher is fully familiar with all the ins-and-outs of the model. As a result, the chance that the newly constructed model meets its specific objective is probably higher than when using existing (supposedly more general) models;
- (ii) *re-use an existing model*: the model, which is supposedly easy to use, is adopted as a whole with only the modification of specific parameter values. However, the researcher often wants to comprehend the whole model (Spitters 1990) instead of viewing the system in terms of its inputs and outputs, without any knowledge of its internal functioning. Furthermore, as the original model was targeted at a specific issue that is most likely different from the one it is re-used for, it may not be optimally suited for the intended application. Moreover, there is a risk of its use outside its validity domain (Monteith, 1996), or of inappropriate parameterisation of the model, as crop models are often over-parameterised (Brun et al., 2002; Chapter 4);
- (iii) *combine both approaches*: this implies developing his/her model starting from an existing one, which is considered, at first, as best suited for the problem to address, and that can be adapted and re-assembled for instance by adding parts of other models. To successfully re-assemble such a new model, intimate knowledge and a thorough understanding of the existing model(s) is required (Argent, 2004). Even for models properly documented, it often

requires intensive interactions with the original crop modeller(s) to decide what has to be changed in the model structure, equations and parameter values for the construction of the new model.

Recently developed crop modelling frameworks tend to support the last option of re-assembly models' parts, highlighting the importance of modularity and flexibility in crop models (Wang et al., 2002; Quesnel et al., 2009; Adam et al., 2010; Donatelli et al., 2010). While these frameworks enable an easy plug-and-play of model components (Papajorgji, 2005) and facilitate modules linkage (Rizolli et al., 2008), it is often difficult for non-experienced users to fully apprehend the tool, to completely understand the different available components and to combine them properly in creating a new model (Chapter 3). Often, the emphasis on the software part of the framework may create a gap between the computer programmer/software specialist and the crop physiologist or agronomist (Argent, 2004; Holzworth et al., 2010), although they both may realize the added value of the use of simulation models for e.g. the formalization of expert knowledge and the design of alternative cropping systems and/or management practices.

The focus of this paper is on the description of a systematic procedure for re-assembly models i.e. developing a new crop model from an existing one, by using a modular crop modelling framework and available crop physiological knowledge. First, the conceptual model, that is seen as a shared "mental model that allows us to understand and simplify the problem" (Fowler, 1997 p.2), is co-developed by scientists of different backgrounds, having theoretical and/or practical knowledge of the crop(s) and cropping system(s) (Wery et al., 2009). Subsequently, this conceptual model of the biophysical system is translated into an operational simulation tool.

Recent progress in crop modelling (Holzworth et al., 2010), and, in a broader context, in modelling for integrated assessment of farm production systems (Ewert et al., 2009), supports this way forward in model building. For example, the crop modelling platform APES (Agricultural Production and Externalities Simulator) is a modular model for cropping systems that can be extended and updated, enabling transfer of new research results to operational tools (Donatelli et al., 2010). It includes components that offer simulation options for different biological processes of relevance to agricultural cropping systems. A component is a piece of software representing plant and/or soil process(es) that is(are) used to compose/construct a cropping system model (e.g. crop, light interception, water uptake, soil water, or soil carbon and nitrogen components). Each component implements the strategy design

pattern (Gamma et al., 1995), enabling easy plug-in of alternative modules, thus providing flexibility to the framework (Adam et al., 2010). A module is a conceptualization of a specific crop or soil process implemented within a component (e.g. radiation use efficiency or photosynthesis for biomass production). Finally, the overarching level of granularity included in APES is the modelling solution (MS), referring to a combination of components and modules to construct one effective simulation model.

The objective of the study is to describe and test a systematic approach to support decision-making on what can be retained from existing models (model structure and parameters) and what needs to be changed (species- or even cultivar-specific modules, equations and parameter values) to adapt it for a new crop system to be simulated. This approach of model re-assembly has been tested using the APES platform, illustrated with a case study in which a grain legume (pea, *Pisum sativum* L.) model is developed from an existing cereal (durum wheat, *Triticum durum* L.) model.

2. Materials and methods

2.1 From one model to another: the need for three disciplines

We re-assemble an existing crop model with new knowledge to simulate performances of a new crop, leading to a new operational crop model (called here modelling solution, MS). The assemblage of the new crop model or MS follows 3 phases, based on 3 scientific domains (Figure 5.1):

- Phase 1: The agronomic domain, consisting of conceptual modelling , to identify the relevant basic crop processes (Wery, 2005), the validity domain of the existing model for system A and the physiological meaning of its key parameters compared to those of the new system (system B);
- Phase 2: The mathematical domain, in which the biophysical processes, identified above, are translated into a coherent set of mathematical equations with associated parameter values;
- Phase 3: The software engineering domain, in which the sequence of the mathematical equations (from phase 2) representing the crop processes (from phase 1) is coded into a computer model. Each set of equations corresponds to a specific conceptualisation of a crop process and is represented by one module.

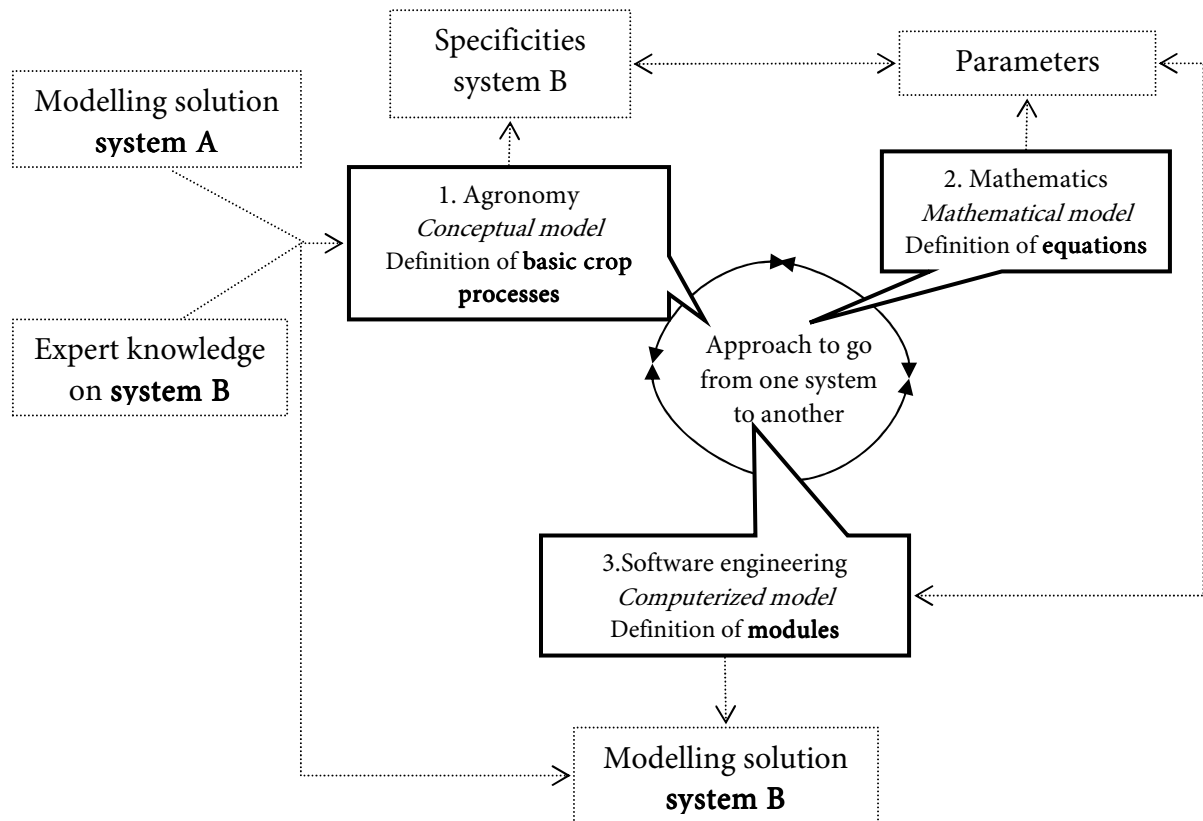


Figure 5.1. Approach to incorporate agronomic expert knowledge of system B (e.g. pea crop) into a new modelling solution from an existing modelling solution of system A (e.g. wheat crop), adapted from Roux et al. (2010).

This approach of model re-assembly is designed to facilitate interactions among the scientific domains, while keeping complex questions specific to a discipline, inside each domain (Roux et al., 2010). Iterative interaction among the 3 domains allows incorporating expert knowledge into the new modelling solution. The use of this approach is based on two main inputs (Figure 5.1): an existing MS (i.e. a crop model for crop A, in the current study the durum wheat model) and expert knowledge of the system to be simulated (crop B, in the current study the pea crop). The main outcome of the approach is a new MS specific to the new crop B.

On the basis of agronomic and crop physiological knowledge, the specificities of the models (e.g. a wheat crop or a pea crop growth model) are formulated: objective of the simulation, boundaries of the system, environment (i.e. forcing functions), key variables and parameters, input and output variables, and relevant relationships between the variables (De Wit, 1968). Such a graphical representation of the system results in a conceptual model, (i) distinguishing the main basic crop processes shared by the two crops from those specific to one of them, and (ii) describing the main hypotheses in a transparent manner (e.g. in an associated table). The simple

representation of the model is a prerequisite to facilitate the exchange between the software specialists and the crop physiologists (Van Daalen et al., 2002). This conceptual model illustrates the hidden and acknowledged assumptions of the system (Heemskerk et al., 2003), identifies the crop generic versus specific processes and equations, and forms the basis for the discussions among the crop specialists. Subsequently, the detailed formulations of the crop processes are included in the mathematical model. The transformation from the conceptual model to the mathematical model introduces a clear structure in the set of equations and yields an explicit description of the differences in the meaning of the key parameters, representing the specificities of system B in comparison to those of system A. Finally, the computerized model is based on the mathematical models by coding the structured set of equations. It should be a consistent transcription of the conceptual agronomic model, with ideally each module of the software representing a basic crop process (Adam et al., 2010).

2.2 Approach to construct a modelling solution for a new crop

The approach builds on the Agricultural Production and Externalities Simulator (APES, Donatelli et al., 2010), a crop modelling framework that contains a set of modules (each conceptualizing a specific crop process) for the simulation of plant growth and development and soil water and nutrient dynamics. Its component-based design makes it flexible and modular, facilitating the development of MSs (i.e. crop models), representing various types of cropping systems. This software architecture enables to capitalize on the new concept without modifying the original crop component, but rather by developing it further (Clements 1995; Argent 2004).

In the approach developed here, modifications in the MS were introduced at three levels: (i) through adapting values of parameters, (ii) within a module, by changing, adding or removing (an) equation(s) with its (their) associated parameter(s), and (iii) in the overall structure of the model, by adding or removing algorithms that simulate (a) specific process(es) (i.e. module(s)). To identify the necessary changes, a clear description of the original MS (i.e. the conceptual model of system A) helped to clarify the hypotheses included in MS_A and acted as an aid in the transfer of knowledge (Sterk et al., 2009). To describe the underlying assumptions in MS_A, we relied essentially on the interaction between the crop physiologist and the crop component developer of APES (i.e. having the knowledge of what is in the existing modules). This interaction allowed an *a posteriori* definition of the basic crop processes, the equations and the software modules underlying MS_A, in a consistent way across the three domains (Figure 5.1).

Table 5.1. Sources of information to collect expert knowledge on the pea crop.

Crop processes considered	Type of expert	Type of exchange (timing)	Main outcome	Source
All	Agronomist and modeller	Workshop (June 2008, October 2008)	Description of the original modelling solution	Shibu et al., 2010, crop component help files (www.apessimulator.it)
N fixation Phenology	Agronomist with general knowledge of crops and more specifically of peas	Workshop (August 2008, January 2010)	Identification of changes needed for the pea crop and qualitative evaluation of the new MS	Wery, 1996
All	Crop physiologists, pea experts (J. Wery, J. Lecoœur and L. Guilioni)	Literature and discussion (beginning 2009)	Data collection Better understanding of pea functioning	Agrophysiologie du pois proteagineux, Munier-Jolain et al., 2005 Lecoœur et al., 1998 Guilioni et al., 2003 Debaeke et al., 2006
Dry matter accumulation	Legume specialists (MH. Jeuffroy, G. Kaschuk)	Discussion (end 2009)	Modelling hypothesis on sink stimulation	Discussion Kaschuk et al., 2009 Yin and Van Laar, 2005

From the conceptual model MS_A and expert knowledge from system B, we identified the processes that could be retained as they were, the ones that needed minor changes (i.e. parameter values, or equations) and the processes that should be added or removed (i.e. as modules), reflecting the specificities of the new system to be modelled. From literature review and data collection, we derived parameter values and modified model equations. This step resulted in an adapted conceptual model, including the main changes needed in the transition from MS_A to MS_B .

The combination of the new modules and the modified equations (including parameter values) forms the new modelling solution (MS_B). The software architecture of the crop component enables to capitalize on these new approaches, while retaining (without modification) the original crop component. Finally, the crop component developer had intensive exchanges with the expert crop physiologist/agronomist to evaluate (qualitatively) the consequences of the changes introduced in the new modelling solution.

2.3 Case study

The methodology described so far, has been applied to transform a (durum) wheat model (MS_{wheat}) into a pea model (MS_{pea}) that can be used to design new legume-based cropping systems in Midi-Pyrénées (Mahmood et al., 2010). The main sources of expert knowledge used are summarized in Table 5.1.

First, to modify the original wheat modelling solution (MS_{wheat}) that is based on the LINTUL3 model (Shibu et al., 2010), we identified three agronomists with substantial general knowledge of crop physiology to get a good understanding of the original wheat modelling solution, and with insights as well in the pea crop. Workshops were conducted with those crop experts and the crop modellers, developing the MS_{pea} . Further, to deepen the insights gained, an extensive literature review and short communications with specialists of legume physiology were carried out.

The main outcome of this exchange of information was a clear conceptual model of the new grain legume MS, with an explicit description of the required modifications in the cereal modelling solution. In the second phase, we translated the specificities of the pea crop into the required modifications at the three levels, i.e. parameter values, equations and algorithms/modules. Having identified these modifications required in the conceptual model for the grain legume, we expressed them into new values for specific parameters (thus becoming crop-specific parameters), new equations (changes in the mathematical model) and/or new modules (changes in the computer model, structure of the tool) to be included in the new MS.

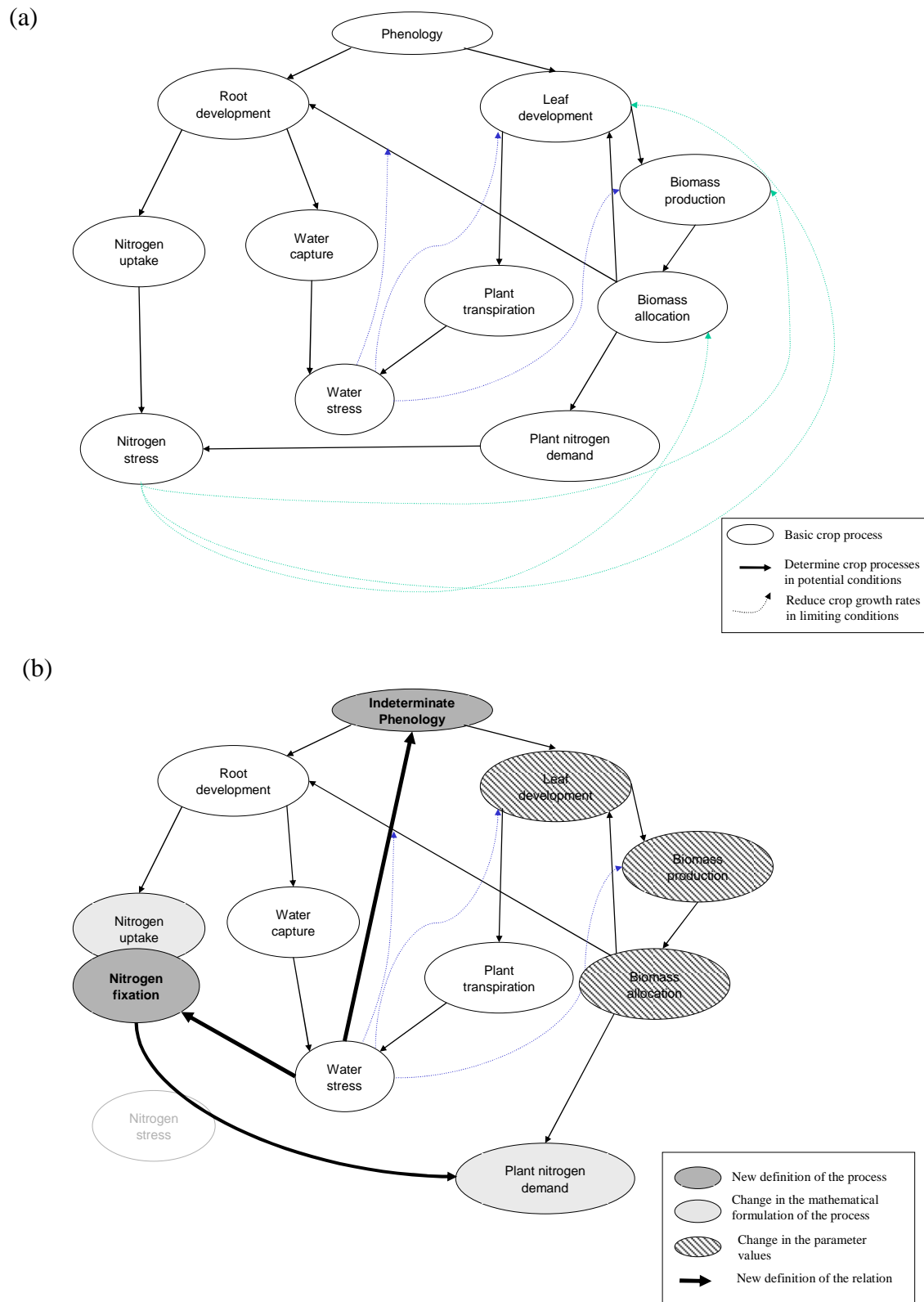
Following iterative cycles, started in mid-2008, with crop physiologists and software engineers, and the translation of the various concepts into mathematical form, the new modelling solution for grain legumes (MS_{pea}) was available for the targeted application by the end of 2009 (Mahmood et al., 2010). A final step before supplying the MS_{pea} to the user was its evaluation. However, the focus of this study being on the development of a new model (i.e. modelling solution), we only validated the new model qualitatively, testing its performance with respect to hypotheses formulated by the crop experts. Since we developed this MS in the frame of a project aiming at designing new cropping systems, based on the introduction of grain legumes in the farming systems of Midi Pyrénées, France (Wery and Ahlawat, 2008, Mahmood et al., 2010), the tests of the model were carried out using data representative for the region of the project. We, therefore, executed a simulation test for the new pea model with weather data from Montpellier from 2003 to 2006 (Zander et al., 2010), and assessed its performance using expert knowledge and data from Wery (1996) and Debaecke et al. (2006). We tested MS_{pea} under potential and under water-limited growth conditions (Van Ittersum and Rabbinge, 1997).

3. Results

3.1. Understanding the wheat modelling solution

The first result of the application of the proposed approach for model re-assembly was a clear description of the original modelling solution of the wheat crop (MS_{wheat}), identifying the main crop growth and development processes and their interactions (Figure 5.2a). In general, crop growth models are developed at different degrees of complexity, depending on the objectives of the study (Penning de Vries, 1982; Sinclair and Seligman, 1996). APES, developed within the SEAMLESS project (Van Ittersum et al., 2008), needed the capability to simulate individual plant organs, because users were not only interested in the grain yield, but also in amount and composition of the plant residues returned to the soil. Overall, the APES model needed the capability to respond to farm management actions in order to generate the so-called technical coefficients for a whole farm model; FSSIM (Farm System Simulator, Louhichi et al., 2010), to be run over a number of years, and, finally, data demands should not be too large. Hence, the crop component included in APES is based on the concept of radiation interception and utilization with allocation of carbohydrates to different plant parts, contained in the LINTUL model (Spitters, 1987): a simpler model would not be able to distinguish organs, whereas a more complex model would need more data, particularly for the description of the photosynthesis process and canopy structure, which are not likely to improve model performance or relevance at the level of the cropping system when field experiment data are limited.

Figure 2a depicts the conceptual model of MS_{wheat} , drawn in intensive interaction between the crop component modeller of APES and experts on crop physiology. It is a simple schematic representation of the different crop processes involved, serving as a basis for exchange with the pea crop physiologists. Crop growth and development can be defined through a few main basic processes (under potential growth conditions) (Wery, 2005): (i) phenological development, (ii) leaf area development, (iii) root development, (iv) biomass production, and (v) its partitioning over the plant organs. Subsequently, as a function of resource availability, growth under resource (water and nitrogen)-limited conditions can be defined through the effects of water and nitrogen (N) stress, quantified on the basis of the balance between water or N demand and uptake (i.e. capture) by the crop.



To share this conceptual model with scientists for its evaluation and to create confidence for users, the basic hypotheses underlying its structure and functions, i.e. “what is [included] in the model and what has been considered as less important” should be made explicit in the description of the model (Wery et al., 2009, p. 286). A hierarchical representation of the various crop processes, derived from expert knowledge, identifies the major components, drivers and functions of the wheat crop. First, phenology drives the development of leaf area and root expansion, both governing the capture of resources (i.e. light, water and nutrients) to produce biomass that is subsequently allocated to the various organs (roots, leaves, stems and storage organs) and to determine the potential transpiration of water. Subsequently, governed by resource availability (i.e. water and N in the soil) and crop demands (i.e. potential transpiration and N demand), crop growth may be limited by water and/or N stress. Table 5.2 synthesizes the main crop processes and the associated representation for each model domain (i.e. agronomy, mathematics and software engineering).

The mathematical model of MS_{wheat} is based on differential equations, relating the rates of changes of the states to themselves, other states and to environmental conditions (Leffelaar, 1999). If the state of a model at a certain moment is known, its state in the future can be calculated by means of numerical integration (the Euler method in this example). Appendix A gives a detailed description of each of the processes included in the crop component for cereals (i.e. wheat) of APES (www.apessimulator.it), including the main equations (i.e. mathematical model).

The computerized model is based on the principle of component-oriented design as implemented in APES, with each process represented in one module (Adam et al., 2010). The crop component of APES uses a structure in which parameters and modelling approaches (i.e. modules) can differ according to the crop simulated. Initially, the crop component has been based on the concept of light interception and utilization. Modifications and additions can be introduced to extend the list of crop types for which the model can be used. These changes include the implementation of alternative modules for each of the main crop physiological processes previously identified: (i) leaf development, (ii) biomass accumulation (Monteith, 1977), (iii) biomass partitioning (Van Keulen and Seligman, 1987), (iv) phenology (Van Keulen and Seligman, 1987; Streck et al., 2003; Hearn, 1994), (v) water dynamics and (vi) N dynamics (Shibu et al., 2010).

Table 5.2. The main basic processes of the wheat modelling solution (MS_{wheat}) and the underlying principles in terms of conceptual, mathematical and computer models.

Crop processes	Conceptual model	Mathematical model *	Computerized model **
Phenology	Photo vernal thermal time: cumulative daily effective temperature (Van Keulen and Seligman, 1987) and delay of the development at the beginning (Stöckle, 2003) by vernalization from emergence until floral initiation and by photoperiodic response from emergence till anthesis	Differential equations and reduction factors	Winter crop phenology module
Leaf development	During the juvenile stage driven by temperature (cell division and extension). After this stage, leaf area expansion is restricted by the supply of dry matter	Differential equations	Leaf area module
Root development	Cellular automaton responding to depth and plant water uptake	Differential equations	Link with the root distribution module
Biomass production	Radiation use efficiency approach (Monteith, 1977; Sinclair and Muchow, 1999)	Linear relation	Biomass production module
Biomass allocation	Biomass is allocated as a function of development stage (Van Keulen and Seligman, 1987)	Distribution factors (allocation table)	Biomass allocation module

Crop processes	Conceptual model	Mathematical model *	Computerized model **
Potential crop transpiration	Penman (1948)	Penman equation	Link with climate module,
Water capture	Depending on potential transpiration, soil water content and soil water characteristics	Functional equations	Link with the water uptake module
Water stress	The ratio between actual (water capture) and potential transpiration decreases linearly with soil moisture suction from unity at the critical suction to zero at wilting point (Feddes et al., 1978). Effect of water stress on 3 main processes (leaf development, biomass production and allocation)	Reduction factor	Crop component wrapper***, through a parameter value (drought tolerance)
Plant nitrogen demand	Difference between maximum and actual organ nitrogen contents	Differential equations	Nitrogen dynamics module
Nitrogen uptake	Partitioning of available soil N between crop and microbial demands	Differential equations	Link with the soil CN component
Nitrogen stress	nitrogen nutrition index approach (Lemaire, 1997) Effect of nitrogen stress on 3 main processes (leaf development, biomass production and allocation)	Reduction factor	Nitrogen dynamics module

* equations are given and explained in Appendix A.

** more information on the source code can be found at www.seamlessassociation.org or www.apessimulator.it.

*** a wrapper is derived from the adapter design pattern (Gamma et al., 1995). It allows classes to work together that normally could not because of incompatible interfaces.

3.2. Modifications in the wheat modelling solution: integration of the expert knowledge in the new modelling solution for pea

Figure 5.2b shows the conceptual model for the new modelling solution (MS_{pea}), containing all the modifications for its re-assembly. This is the main outcome from the interactions and exchanges (workshops and discussions) with experts on legumes, and more specifically on the pea crop, complemented with knowledge derived from a literature review. On one hand, some of the modifications in the conceptual model can result in changes in parameter values or in a new mathematical definition of the processes in the model (Table 5.3, change in mathematical model). On the other hand, at the software level, the definition of new processes can lead to the addition of new modules in the computerized model or in the removal of existing modules (i.e. Table 5.3, change in the computerized model).

We identified the modifications needed to construct MS_{pea} as follows (Figure 5.2b):

- (i) Changes in parameter values: different values for specific leaf area (SLA), radiation use efficiency (RUE) (Kaschuk et al., 2009), and modification of carbohydrate allocation as a function of development stage;
- (ii) Changes within a module: addition of an equation to limit N uptake from the soil (Wery, 1996) and;
- (iii) Changes in the overall structure of the model: addition of a N fixation module (Wery, 1996) and replacement of the phenology module (called indeterminate phenology) to define the indeterminate pattern of the plant, i.e. an extended flowering period under the control of temperature and water stress (Ney et al., 1994).

(i) Changes in parameter values correspond to “common” changes in any crop model to adapt it to a specific species (Boogaard et al., 1998). Three parameters have been identified as having significantly different values for the pea crop compared to wheat (Figure 5.2b). Specific leaf area (SLA) defines how much leaf area is associated with a unit (increment in) leaf dry weight. Observations on wheat and pea leaf (including tendrils) morphology suggest that SLA of the two crops is different. Pea leaves are thinner than those of wheat and consequently, their SLA is higher (0.022 *vs* 0.018 m² g⁻¹). From literature review and data collection (Debaeke et al., 2006), we estimated that SLA of pea ranged from 0.033 (Yin and Van Laar, 2005) to 0.025 m² g⁻¹ for various cultivars (from semi-leafless to leafy).

Table 5.3. Summary of the main changes needed for MS_{pea} as compared to MS_{wheat} according to the disciplines.

Main crop growth and development processes	Change in the conceptual model	Change of mathematical model Change within module: equation with associated parameter	Change in the computerized model/the software design Change in the structure
Phenology	No vernalization, no photoperiod, addition of a parameter for the shorten flowering time flowering period due to water stress,	NC*	Need for a new module
Leaf development	Pea leaves are thinner than wheat leaves	Change of the parameter <i>specific leaf area</i>	NC
Root development			NC
Biomass production	Protein production is more energy-demanding than carbohydrate production No additional carbon cost due to N fixation	Change of the parameter <i>radiation use efficiency</i>	NC
Biomass allocation	Allocation of biomass to the grain starts before anthesis	Change of parameter value in the <i>allocation table</i>	
Crop transpiration			NC
Water capture			NC
Water stress	The contribution of N_2 fixation is reduced and flowering period is shortened under drought conditions	NC	Change in model structure associated with the N capture module and phenology module
Crop nitrogen demand	Nitrogen demand to be satisfied by the soil is limited due to additional contribution from nitrogen fixation	Addition of equation with associated parameter	NC
Nitrogen uptake	N fixation strategy complementary to the N-uptake module	NC	Need of a new module
Nitrogen stress	No stress	NC	Remove the nitrogen nutrition index module

* NC: no change required.

A similar reasoning applies to radiation use efficiency (RUE), the slope of the linear regression between accumulated intercepted photosynthetically active radiation (PAR), and accumulated total biomass (Monteith, 1977). It summarizes the combined effect of photosynthesis and respiration processes. Experimental evidence convincingly shows that this parameter is crop-specific (Sinclair and Muchow, 1999). As protein production is more energy-demanding than carbohydrate production (Penning de Vries, 1974), RUE should be lower for legumes than for cereals. A value of 2 g dry matter (DM) $(MJPAR)^{-1}$ was estimated for pea from data collected by Debaeke et al. (2006), in agreement with recent literature (Giunta et al., 2009), compared to a value of 3 g DM $(MJPAR)^{-1}$ for wheat.

Furthermore, Kaschuk et al. (2009) indicate that the carbon costs of the rhizobial and arbuscular mycorrhizal symbioses for legumes are compensated by increases in the rates of leaf photosynthesis. Hence, no additional mechanisms need to be modelled to represent the cost of N_2 fixation on assimilate production, as is generally done in more complex models (e.g. GECROS, Genotype-by-Environment CROp growth Simulator, Yin and Van Laar, 2005). Finally, values for the coefficients of allocation of dry matter to the different organs were modified from the wheat to the pea crop according to the information from Munier-Jolain et al. (2005, p.115). Although the values for the N concentration in the different organs were also modified, we did not consider it as changes in ‘parameter values’, since it also implied the inclusion of new equations in the mathematical model (see (ii)).

(ii) We identified major changes in the equations included in two modules, i.e. N demand and N uptake (Figure 5.2b). We assumed that pea has a lower potential of N uptake than wheat, which may be related to a lower fine root density in the surface soil layers (Gregory, 1998) and a lower “activity of nitrate reductase” (ANR, Wery, 1996). Therefore, we adjusted the N demand of the pea crop to define the required N uptake from the soil, by a reduction coefficient k (Figure 5.3a), meaning that, for a given soil and N demand by the shoots, pea will absorb less mineral N from the soil than wheat. Crop N uptake is calculated from the nitrate N available in the soil. Daily N fixation from the atmosphere is then defined as the difference between daily crop N demand and daily N uptake from the soil (Figure 5.3a), assuming absence of N stress in the pea crop (see below).

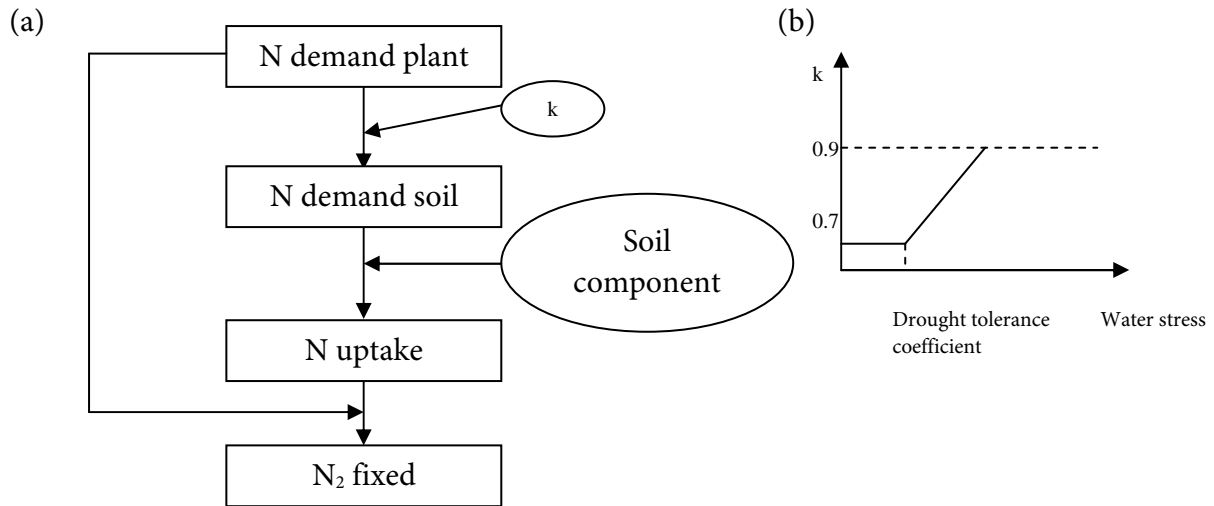


Figure 5.3. Nitrogen (N) demand and N uptake, i.e. N uptake from the soil and N₂ fixation, (a) conceptual approach for the new modules to be included in MS_{peas} (b) with the reduction factor (k) for N uptake from the soil dependent on water stress.

(iii) From the interaction with the crop experts, we identified two main modules to be added (i.e. N fixation and indeterminate phenology, Figure 5.2b), and one to be removed (N stress, Figure 5.2b). In formulating the N fixation module, we assume that N₂ fixation can fully satisfy the N requirements of the legumes during the growing season, if N uptake from the soil is insufficient (Wery et al., 1986; Voisin et al., 2007). Although N deficiency can be observed in pea (Munier-Jolain et al., 2005), we did not consider that phenomenon, as it is often the result of phosphorus and/or potassium deficiency or pea leaf weevil (*Sitona lineatus*) infestation (Wery, 1996), factors not considered in our initial MS. Further, we assumed that no N fixation takes place before development stage 0.3, the time needed for establishment of the nodules during the early phase of growth (Munier-Jolain et al., 2005). The contribution of N fixation to N uptake is assumed to decrease as water stress increases (Mahieu et al., 2009), translated by the increase in the reduction coefficient k with water stress (Figure 5.3b). This phenomena can be linked to the accumulation of nitrates in leaves to sustain osmotic adjustment (Lecoeur et al., 1992).

In formulating the indeterminate phenology module, the indeterminate behaviour of the crop was defined to be governed by the temperature regime, but is co-determined by water stress. As in MS_{wheat} , crop development, i.e. the order and rate of appearance of vegetative and reproductive organs, is defined in terms of phenological developmental stage (DVS) as a function of temperature sum, i.e. cumulative daily effective temperature. However, neither vernalization nor photoperiodic response was included. Further, we defined the length of the flowering period in dependence of water stress. The flowering period is shortened under water stress conditions (Ney et

al., 1994). Guilioni et al. (2003) demonstrated that under moderate water stress the number of flowering nodes is reduced. Hence, we defined the end of flowering in terms of a temperature sum, but also as a function of a water stress index (Figure 5.4).

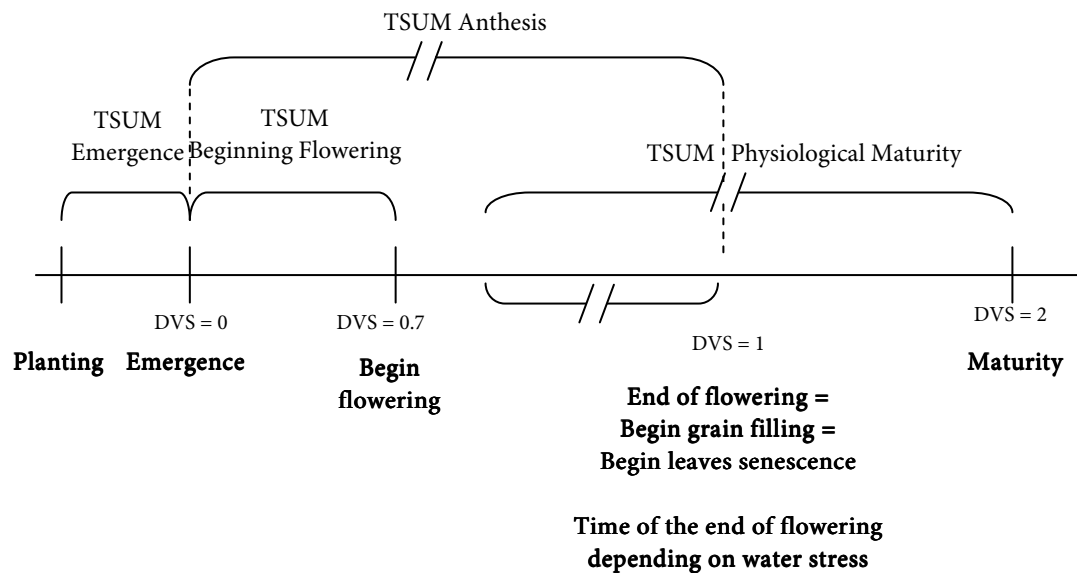


Figure 5.4. Indeterminate phenology module for MS_{pea} : representation of the development stages co-determined by temperature (TSUM) and water stress.

Appendix B gives the mathematical description of the new modules included in the crop component for legumes (i.e. MS_{pea}) of APES (www.apessimulator.it). Each new module has been implemented as an alternative module for the corresponding crop physiological processes in the computerized model, based on the principle of component-oriented design (Adam et al., 2010). We included a new module called “indeterminate phenology” under the phenology physiological process and a module called “N fixation” under the N dynamics physiological process. The final MS_{pea} includes the following modules: (i) leaf development based on temperature and leaf dry matter production, (ii) biomass accumulation based on *RUE* (Monteith, 1977), (iii) biomass partitioning based on allocation factors (Van Keulen and Seligman, 1987), (iv) indeterminate phenology, (v) water dynamics and (vi) N dynamics (Shibu et al., 2010), including a N fixation module. Table 5.4 gives the key parameter values for MS_{pea} compared to MS_{wheat} .

Table 5.4. Key parameters for both modelling solutions: MS_{wheat} and MS_{pea} .

Modules	Parameters	Unit	Default values	
			MS_{wheat}	MS_{pea}
Phenology	Air temperature (base)	°C	0	0
	Air Temperature (Sum Emergence)	°C d	160	140
	Air Temperature (Sum Anthesis)	°C d	550	1200
	Air Temperature (Sum Maturity)	°C d	840	700
	Maximum number of days for vernalization	d	50	n.a.
Additional parameters for the indeterminate phenology module	Air Temperature (Sum Beginning Flowering to Maturity)	°C d	n.a.	700
	Water Stress delaying flowering	unitless	n.a.	0.2
Leaf area development	Leaf Area Index (Relative Growth Rate Exponential Phase)	(°C d) ⁻¹	0.005	0.007
	Specific Leaf Area	m ² g ⁻¹	0.020	0.025
Biomass production and partitioning	Radiation Use Efficiency	g MJ ⁻¹	3	2
	Fraction of dry matter reallocated from leaves to the grains	-	0.4	0.3
	Fraction of dry matter reallocated from stems to the grains	-	0.2	0.2
Nitrogen dynamics	N (Maximum Concentration Storage Organs)	g N m ⁻²	0.025	0.06
	N (Maximum Concentration Leaves at initial conditions)	g N m ⁻²	0.06	0.08
	Fraction (Maximum N Concentration Root of N Concentration Leaves)	-	0.37	0.37
	Fraction (Maximum N Concentration Stem of N Concentration Leaves)	-	0.40	0.5
	Translocation (Time Coefficient of N)	d	10	10
Additional parameters for the nitrogen fixation module	N uptake reduction factor	unitless	n.a.	0.7-0.9
	Drought tolerance coefficient	unitless	n.a.	0.2

3.3. Outcome from the pea modelling solution

We present the simulation outputs of the new pea model (MS_{pea}) to test a number of hypotheses derived from the workshop conducted with the crop physiologists specialized in pea. We first assessed the performance of MS_{pea} with respect to N dynamics under potential growth conditions. It was assumed that N fixation progressively replaces N uptake from the soil to satisfy crop N demand (Deschamps and Wery, 1987). Figure 5a shows such a pattern; we observe three main phases in the dynamics of N fixation and crop uptake of soil N (Figure 5.5b) related to crop growth (Voisin et al., 2007): (i) nodule establishment, when only soil N uptake takes place, (ii) vegetative growth phase, where daily N fixation is correlated with biomass growth (Hooda et al., 1989) and exceeds soil N uptake, and (iii) the grain filling period, when N fixation gradually declines with the decrease in crop N demand (Wery et al., 1988).

Secondly, we assessed at the main differences in pea productivity under different water availability regimes (i.e. intensity and timing), to test three hypotheses on the impact of water stress on pea growth and development, formulated on the basis of the discussions with experts and corroborated by literature data:

- (i) The end of flowering is governed by the temperature regime, but is co-determined by the degree of water stress (Ney et al., 1994);
- (ii) The contribution of N_2 fixation to total N-accumulation in the crop is lower under drought conditions and under conditions of high soil N supply (Zahran, 1999; Mahieu et al., 2009);
- (iii) A higher grain yield under moderate water deficit, compared to well-watered crops, is reproduced by the model as expected from the literature (Wery, 2005).

Table 5.5 summarizes the main outputs of the simulations with different water limitations.

- (i) Moderate water stress during the flowering period (i.e. constant low water stress applied as a forcing function) results in shortening of the flowering period by nine days (corresponding to about 200 degree-days, Table 5.5), due to formation of a reduced number of flowering nodes (Guilioni et al., 2003). When the stress mostly occurred after the flowering period, no major reduction in its length was observed.

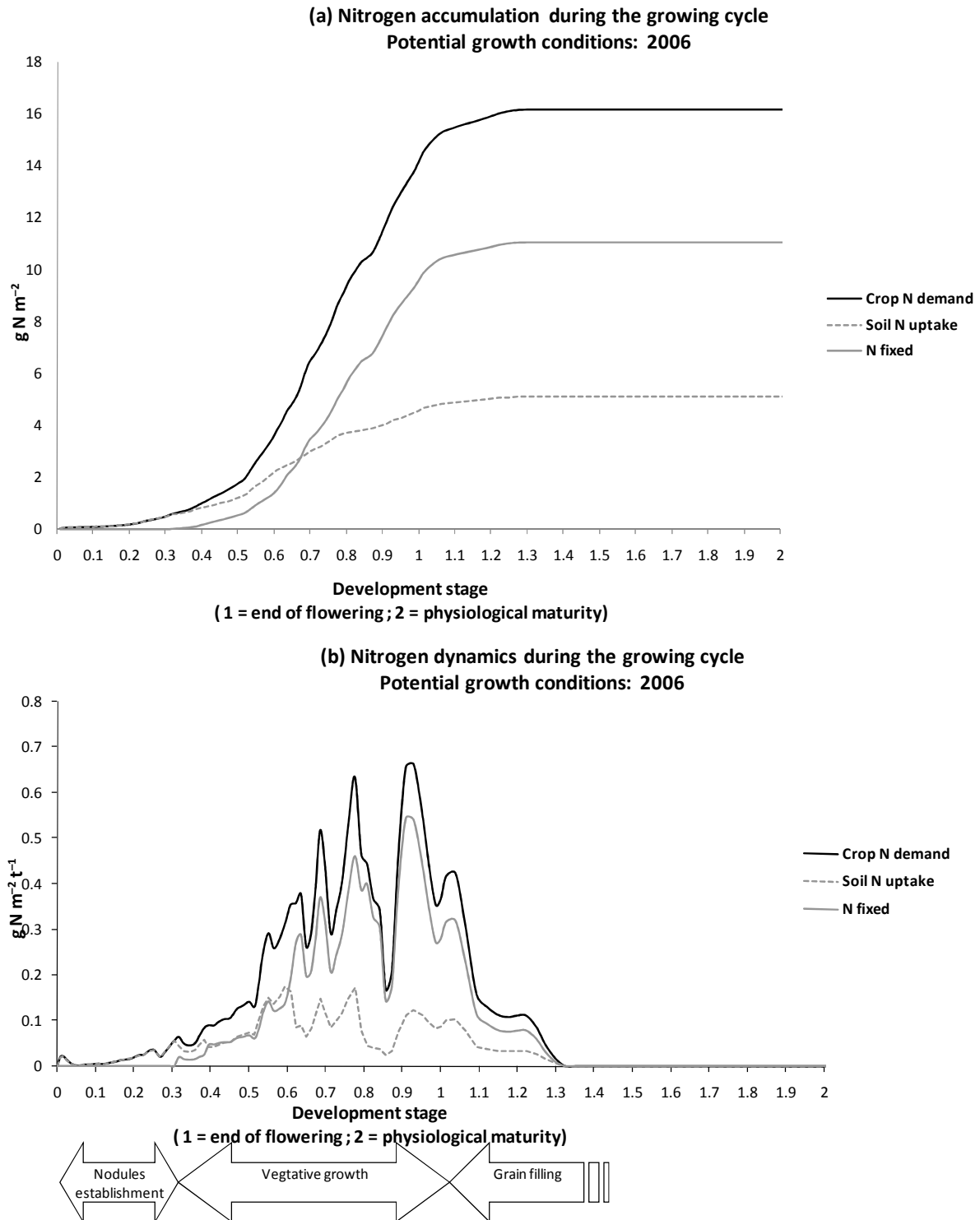


Figure 5.5. Nitrogen dynamics in pea under potential growth conditions*:

(a) accumulated crop nitrogen demand, nitrogen uptake from the soil and nitrogen fixation over the growing cycle;

(b) daily crop nitrogen demand, nitrogen uptake from the soil and nitrogen fixation over the growing cycle. Integrating Figure 5b gives Figure 5a.

* during the 2006 growing season no water and nitrogen stress occurred and there were no nitrogen losses through leaching.

Table 5.5. First simulation results to assess that MS_{pea} produces “reasonable” outputs. Evaluation of the output was done during a workshop involving a pea crop expert, and with data from the literature.

Scenario *	Length of flowering period d	Aboveground biomass t ha ⁻¹	Yield t ha ⁻¹	Soil nitrogen uptake kg ha ⁻¹	Nitrogen fixed kg ha ⁻¹	Contribution N fixed to total N accumulation over the crop cycle %
Potential growth	27 (±0.4)	13.2 (±0.8)	4.6 (±0.3)	51(±24)	106(±20)	67
Moderate water stress during flowering period	18 (±0.4)	9.3 (±0.9)	5.0 (±0.6)	34 (±13)	54 (±13)	61
Severe water stress from flowering to physiological maturity	26 (±0.4)	9.4 (±0.6)	2.1 (±0.2)	51 (±24)	95 (±14)	64

* average values over years 2003-2006; standard deviation in parenthesis

(ii) Data from Wery (1996) suggest that in a N-rich environment, N fixation contributed about 70% to total N accumulation over the crop cycle under potential growth conditions, compared to 60% under water-limiting conditions, where N uptake from the soil compensated the reduction in N fixation under water stress. Our results are of the same order of magnitude (Table 5.5), confirming a lower contribution of N_2 fixation to total N-accumulation in the crop under drought conditions and under conditions of high soil-N supply (Zahran, 1999; Mahieu et al., 2009).

(iii) A higher grain yield under moderate water deficit, compared to potential growth conditions, is reproduced by the model (Table 5.5), as observed by Wery (2005) for indeterminate legumes such as pea and chickpea, and for field pea by Turc et al. (1990), while aboveground biomass was about 20 to 40% lower (Turc et al., 1990; Munier-Jolain et al., 2005). This behaviour is attributed to the fact that in the absence of drought (i.e. under potential growth conditions) excessive leaf growth results in increased flower and pod abortion due to competition for carbohydrates, while moderate water deficit reduces leaf expansion more strongly than photosynthesis (Wery, 2005). Although these physiological processes are not explicitly represented in the model, it shows the emergent property of higher grain yield under moderate water deficit (Table 5.5). On the other hand, severe water stress from flowering to physiological maturity, a common phenomenon in Mediterranean regions (e.g. Montpellier), results in substantially lower grain yields (Mahieu et al. 2009; Table 5.5).

4. Discussion

4.1 Roles of the different models/disciplines

The methodology of model re-assembly illustrated in this study, shows that construction of a crop growth model requires different disciplines, from crop physiology to computer science, with intensive interaction among them. Each of the disciplines plays a specific role in each of the phases of the crop model (re)assembly (from a conceptual model to a software tool). First, the (physiological) conceptual model can be seen as an “eye opener” (Van Daalen et al., 2002), identifying the main characteristics of the system under study and the objective(s) of the simulations. The development of this conceptual model relies on expert knowledge, combined with modelling skills, using principles of system analysis (De Wit, 1968; Odum, 1983). Subsequently, through a graphical representation, the explicit description of the validity domain and the main interpretations of the key parameters of the system, the conceptual and mathematical models can act as “arguments in dissent” or “vehicles in

creating consensus” (Van Daalen et al., 2002). Finally, the computerized model is often used as a software tool for designing and/or evaluating “management options” in research or development programs. It may support the user in making choices on the best agronomic management practices (i.e. an operational tool) for a given set of objectives. For instance, APES was initially designed within the SEAMLESS Integrated Framework to simulate the biophysical behaviour of a wide range of crops in specific rotations in response to agro-technical management options in interaction with weather and soil conditions (Donatelli et al., 2010). As a concrete example, in our case study, we could demonstrate that moderate water stress during the flowering period resulted in higher pea grain yields and a reduction of vegetative biomass and nitrogen fixation.

To represent the conceptual model, we decided to use neither a Unified Modelling Language (UML) nor a relational diagram such as a Forrester diagram (Forrester, 1961). Donatelli et al. (2002, p.8) identified the use of UML or any “visual tools” as a promising way forward for crop modelling. We support this conclusion on the use of visual tools to conceptualize the model and as a means to promote transparency in crop modelling. However, we believe that use of these types of diagrams in our approach may bias the discussions and model-re-assembly exercise in favour of modellers and/or software specialists. UML is a standardized general-purpose modelling language in the field of software engineering, while the relational diagram originated in the system analysis field. Using a rather coarse representation removed the “bias” towards a specific discipline and did facilitate the interaction among the 3 scientific domains.

4.2. Re-assembly from “building stones”

Our approach for model re-assembly uses new software techniques that support modularity and flexibility of the modelling tool. It can be seen as an efficient way to create *ad hoc* models, rather than re-inventing the wheel (i.e. creating new models from “scratch”). Reuse of existing models that are supposedly seen as generic often leads to development of a large number of “scattered” versions of such models, with poor traceability of changes in parameters and equations with regard to crop physiology knowledge. The original model is often tailored to a specific objective, which determines to a large extent its structure and level of detail and complexity (Brooks and Tobias, 1996). Although efforts have been made to develop common model algorithms and structure to simulate a wide range of crop species (Boogaard et al., 1998; Brisson et al., 2003), models often need to be adapted for a new situation. The “universal model” that can be applied to any situation does not exist, and

modellers always face the dilemma of either creating their own model or adapting an existing one (Boote et al., 1996).

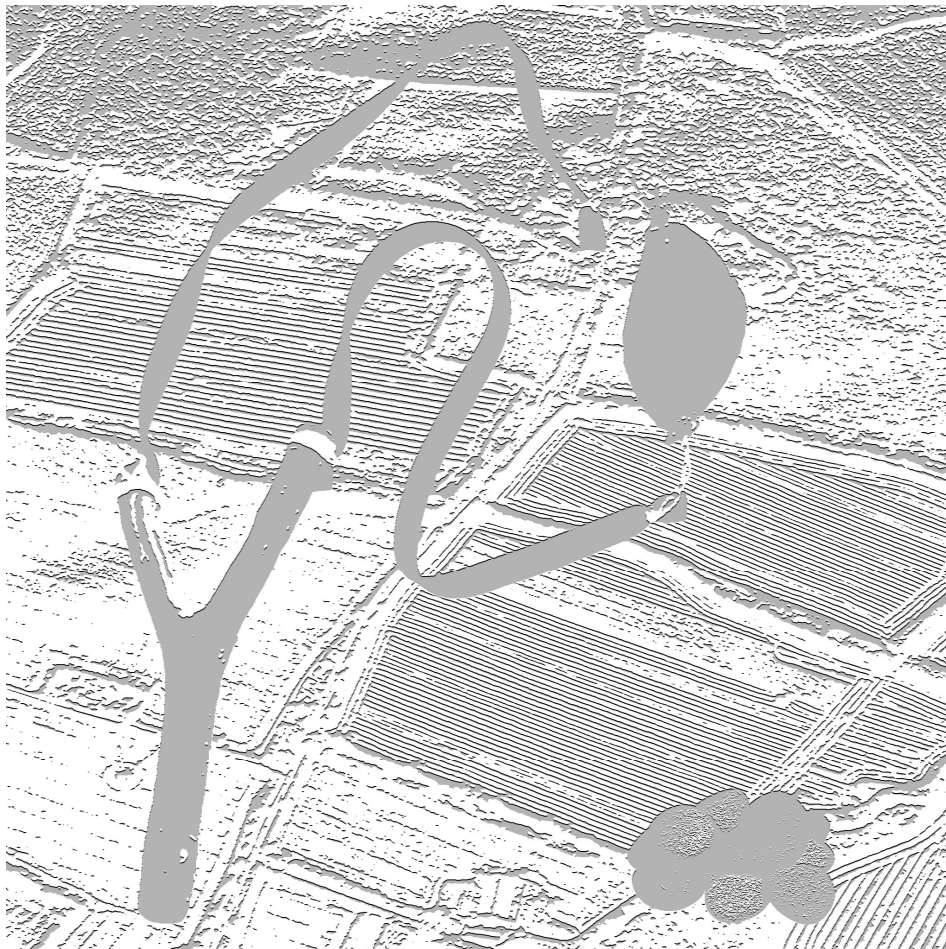
With our methodology we demonstrated that new crop models can be created from “building stones”, using an existing modelling solution as a basis for the creation of a new one, without relying on the model developer. Building on existing models was already advocated by Spitters in 1990 (p. 391), in stating that “the appropriate model version can be composed from [*existing*] building-stones”. Today, new software techniques provide the means to operationalise this principle. The methodology we present is an addition to the existing APES framework in order to facilitate an accurate combination of the “building stones” or modules to meet the modelling objectives and with the available data, in strong interaction with crop physiologists not necessarily experts in crop modelling.

4.3. Approach to combine with modular crop modelling frameworks

We used APES as our support modelling framework. The flexibility and modularity of APES simplify the process of plug-in plug-out of crop or soil processes (Papajorgji, 2005). However, our methodology on model re-assembly should be seen in a more generic sense and could be applied within other existing crop modelling frameworks such as APSIM (Agricultural Production Systems sIMulator, Wang et al., 2002) or RECORD (“REnovation et COoRDination de la modélisation de cultures pour la gestion des agroécosystèmes”, Quesnel et al., 2009). For instance, with respect to APSIM, there are similarities between the use of our approach with the crop component of APES and the use of the generic PLANT model (Wang et al., 2002). “The PLANT model consists of many classes that have been designed in such a way to facilitate their swapping in and out for different crops. [...] the model developer is able to construct new crop models entirely from configuration documents.” (p. 889, Holzworth and Huth, 2009). This easy plug-in plug-out present in these crop modelling framework assumes that these processes can be combined without any problem for the re-assembly of a new crop model (e.g. through new configuration files). However, there is a need to couple this easy plug-and play of modules with a “decision tree” to facilitate the development of these configurations files. The approach we illustrated in this study aimed at facilitating the construction of these files (i.e. selection of the different modules to combine), especially with respect to the criteria of crop type. It helps to combine software architecture with decision criteria, defined by the crop expert, to construct/re-assemble a new sound crop model.

CHAPTER 6

GENERAL DISCUSSION



1. Introduction: combining software techniques with systems analysis

The general objective of this thesis was to develop a modelling framework (i.e. an operational tool and guidelines that support its use) to design and assess innovative cropping systems. The key idea was the development of a modular approach that allows construction of modelling solutions (crop growth simulation models) aiming at answering specific agro-ecological questions through combination of models (or parts of models), guided by the demand for modelling detail and output. Adoption of advanced software engineering techniques in crop modelling in the past decade have led to the construction of modelling frameworks, consisting of libraries of models from which selections can be made, governed by the objectives of a specific simulation exercise. Advantages of a modular structure include the possibilities for: (i) interchange of code among models, (ii) testing alternative hypotheses, (iii) use of simple or comprehensive modules as required for a particular application, and (iv) sharing of expertise.

Although these advantages are undeniable (Acock and Reynolds, 1989), and were illustrated on a few occasions, mostly within the APSIM¹ framework (e.g. Van Oosterom et al., 2006, Moore et al., 2007, McMaster and Hargreaves, 2009), to my knowledge, no research has explicitly addressed the process of module comparisons or model adaptation within such frameworks. The focus is often more on the outcomes of the overall model than on the description of how models are constructed and assembled from different modules. However, as modelling frameworks provide technical possibilities to link modules, even if these links are physiologically meaningless, there is a need to support the qualified selection of modules to be linked. Thus, the main contribution of this thesis is on the model building process, and more specifically on the decision-making process of selecting one module rather than another (modelling beam of the framework) and incorporating that module into the model structure (i.e. module assembly). This selection process is based on explicit criteria to guide model development. Those criteria are identified in this thesis.

In this thesis, I identified three main steps in the development of a framework to introduce flexibility in crop modelling:

- Create the building blocks: identify the relevant library parts and include these building blocks into the framework (Chapter 2);
- Link the building blocks: technically (Chapter 2) and conceptually (Chapters 3, 4 and 5);
- Evaluate and apply the framework (Chapters 3, 4 and 5).

¹ APSIM: Agricultural Production Systems sIMulator (Keating et al., 2003).

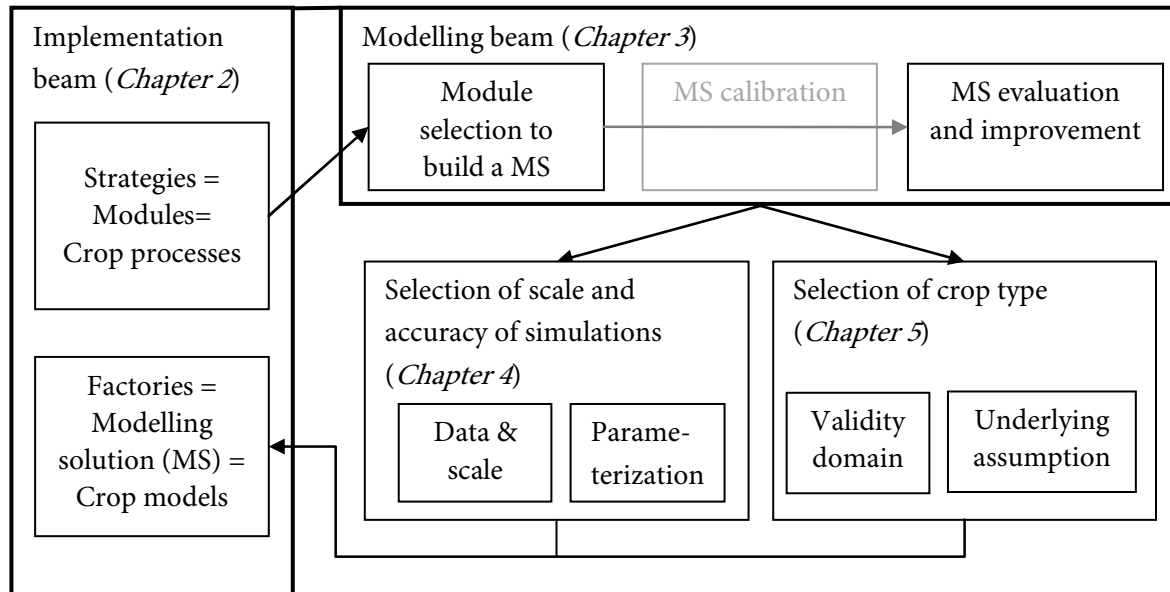


Figure 6.1. The contribution of this thesis to the development of a crop modelling framework consists of (i) two main libraries: the library of processes, and the library of models (Chapter 2), and (ii) concepts: result of the definition of logic for model assembly (Chapters 3, 4 and 5).

Within this thesis (Figure 6.1), I created the building blocks (Chapter 2) and developed methodological approaches (Chapters 3 and 5) to evaluate the crop model structure (i.e. modelling solution, MS) with emphasis on the process of selection of modules (i.e. building blocks) and their linkage (Chapter 4 and 5). The aim of this chapter is to discuss the choices made in the development of the crop modelling framework at the implementation (Section 2.1) and modelling levels (Section 2.2), and to discuss how such a framework can contribute to the advancement of crop modelling, but also its limitations (Section 3). Finally, this chapter concludes with the main contributions of this thesis to the crop modelling field (Section 4).

2. The framework

2.1. Choice of software design and level of granularity: implications for the user

CROSPAL² (Chapter 2), APES³ and APSIM are examples of modular frameworks, illustrating how modularity has been applied so far in crop modelling. Reflections on differences and similarities of the software design adopted to build modular frameworks (i.e. implementation beam) should help in identifying what is essential to

² CROSPAL: CROp Simulator: Picking and Assembling Libraries (Adam et al., 2010).

³ APES: Agricultural Production and Externalities Simulator (Donatelli et al., 2010).

create the libraries, without creating “Yet Another Modelling Framework” (Van Evert et al., 2005) and which are the consequences of their differences.

The CROSPAL software design (Chapter 2) builds on my initial involvement in the development of APES. The shared use of the strategy design pattern⁴ to create the library of basic crop processes (Table 6.1) is one of the outcomes of these activities. The use of the strategy design pattern enables implementation of alternative modules to simulate the same (crop growth and development) process. I used the same strategies for the crop component of APES as the ones described in CROSPAL (based on the basic crop growth processes, Table 2.1, Chapter 2). This definition of modules, i.e. strategy, according to crop growth and development processes was similarly adopted within the generic crop model template (GCROP) of APSIM (Wang et al., 2002), where it is represented by a crop process library (CPL). In a CPL, such processes are compiled in separate dynamic link libraries (dlls) that can be called by the generic model structure. Use of the strategy design pattern, as well as of separate dlls, enables easy addition of new modelling approaches (i.e. modules). It creates a high modularity in the framework and represents the characteristics of crop growth and development through their delineation in basic processes (Wery, 2005): this delineation is similar in all three frameworks (see also Table 2.1. in Chapter 2: comparison of CROSPAL strategies with GCROP delineation).

The main difference among the three frameworks resides in how the overall structure of the crop model (i.e. assembly of modules) is configured. In CROSPAL, the overall structure is obtained via the use of abstract factories and a graphical user interface (GUI), to relate a concrete factory⁵ to criteria of selection (Table 6.1). The use of the abstract factories design pattern (Chapter 2) enables the definition of the basic structure of the model, and, through formulation of new concrete factories, facilitates the creation of a new structure of crop models (i.e. modelling solutions). Further, the use of an abstract class⁶ provides the flexibility for the future user to include physiological principles common across crop types, to the “abstract crop”, providing generality to the model structure (i.e. identification of similarities among crops or generic crop characteristics).

In APES, and more specifically in its crop component, the assembly of the various basic crop processes is constructed via the use of the interface⁷ IStrategy (Table 6.1).

⁴ Strategy design pattern: define a family of algorithms, encapsulate each one, and make them interchangeable.

⁵ Concrete factory = modelling solution.

⁶ Abstract class: a set of operations which all objects that implement the protocol must support.

⁷ Interface: a set of named operations that can be invoked by users.

“A composite strategy [IStrategy] differs from a simple strategy, because it needs other (simple) strategies to provide its output(s)”, (Donatelli et al., 2010, p. 89). In other words, this composite strategy (Istrategy) defines the model structure by invoking other classes⁸ defined as simple strategies. This composite strategy (defining the model structure) can be selected by the user through the selection of a “model option”, either via an XML configuration file (in the integrated version of APES) or via a graphical user interface (GUI) in the stand-alone version. However, in contrast with CROSPAL, neither an explicit link is made via the GUI to relate a specific composite strategy to an MS, nor is a common behaviour to express generality in crop models (similarities among crops) included in the composite strategy.

In the generic PLANT model (Holzworth and Huth, 2009), the modules are combined directly via XML⁹ files (Table 6.1), enabling the basic crop processes to be turned on and off (i.e. calling or not the different dlls). Therefore, the user can define the structure of the crop model him/herself, with no pre-packaged solution as in CROSPAL or APES. The use of XML files in the generic PLANT model enables to completely externalize the configuration of the crop model structure and to simplify the reuse of models (Holzworth et al., 2010).

Table 6.1. Comparison of the different software designs adopted in different crop modelling frameworks, with respect to the level of granularity. (GUI, graphical user interface; XML, Extensible Markup Language)

Level of granularity	CROSPAL	APES	GCROP
Modules <i>Basic crop processes</i>	Strategy	Strategy	Dynamic link libraries (dlls)
Component <i>Crop</i>	Abstract factory and criteria with a GUI	Composite strategy (IStrategy: interface)	Generic model structure/ XML configuration
Modelling solution <i>Soil-crop</i> (i.e. crop simulator)	Definition of new concrete factories (further research needed to include the soil, see discussion of Chapter 2)	Components linked via wrapper (using the ModCom framework : Hillyer et al., 2003)	GCROP linked to the APSIM engine

⁸ Class, here, is used in relation to computer science.

⁹ XML: Extensible Markup Language.

The software design of all three frameworks discussed allows modularity and flexibility in adapting the model structure. But the question then is which degree of flexibility should be given to the user of the framework? I believe that it should vary depending on the expertise of the user. The choice for one specific design mostly depends on how “free” and “knowledgeable” the future user of the modelling framework will be.

Use of an XML file to configure a model (i.e. define its structure) provides total freedom to the user to select any module (Moore et al., 2007), no matter whether the different modules “fit” together conceptually. Technically, everything is possible and modifications in the configuration of the structure of the model are completely externalized. In APES, the use of the composite strategy provides less freedom to the user, as the developer defines this composition within the component (IStrategy) on the basis of his own opinion on the anticipated future modelling exercise/application. However, the selection of a specific model structure still remains the responsibility of the user through the use of model options. Further, the design adopted in APES also enables easy extension of any component and redefinition of the composite strategy (Donatelli et al., 2010) that could be tested by the user. In CROSPAL, the choice for the use of the abstract factory¹⁰ relies on the logic to assemble the crop model. This logic is the consequence of the vision of the developer on crop functioning and should correspond to the different criteria included in the graphical user interface. The selection of a specific MS (i.e. a specific model structure for an application) is guided by the selection of criteria by the user through the graphical user interface (Figure 2.1, Chapter 2).

Finally, both APES and GCROP operate at a higher integration level in the delineation of the system, defining cropping system (soil-crop, water and nutrient limited production), whereas CROSPAL defines the crop system (crop only, potential production). In APES, the crop and soil components are linked using wrappers¹¹, facilitated by the use of MODCOM, a software framework to assemble simulation models (Hillyer et al., 2003). In APSIM, the crop and the soil components are connected through the APSIM engine by receiving and processing ‘messages’ (Wang et al., 2002). A message is defined as an instruction for a component to perform an action. With respect to CROSPAL, further development is required to include the soil component in the framework (see discussion Chapter 2). Incorporation of the soil part could be done through the formulation of new concrete factories and by including new strategies.

¹⁰ Abstract factory patterns: a way to encapsulate a group of individual factories that have a common theme

¹¹ Wrappers are classes enabling combination of other classes that could not be combined, because of incompatible interfaces (i.e. adapter design pattern, Gamma et al., 1995).

2.2 From modules to an MS: a concept guiding the use of the framework

In Chapter 1, I defined the word concept as the “logic to assemble the appropriate modules” and stated that a concept can be constructed on the basis of criteria for module selection or of a systematic approach. The definition of this logic is the main focus of a significant part of this thesis (Chapters 3, 4 and 5), testing different model structures according to the criteria defined in the CROSPAL graphical user interface (Figure 2.1, Chapter 2). The presence of the soil-crop system in APES, compared to the single crop system in CROSPAL, enabled further research on the relation between criteria and the resulting model structure corresponding to the simulation objective. As I was a member of the team that developed APES, it was possible to use this framework to test the different hypotheses on model structure (Chapters 3 and 5), directly working on the code.

I used three main approaches to construct these concepts (Figure 6.2): (i) uncertainty matrix, (ii) sensitivity analysis, and (iii) expert elicitation. Use of these approaches highlighted the importance of the modelling beam of the framework (i.e. the protocol to guide users in the selection of modules). Particularly, emphasis is on the uncertainty assessment of the derived model structure. Without this modelling beam, there is a risk of misuse of the implementation beam of the framework, when the assumptions, underlying the building blocks (i.e. modules), have not been explicitly defined.

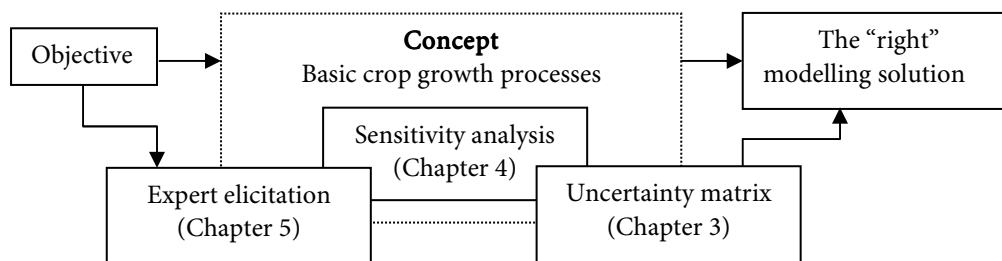


Figure 6.2. Approaches used (in this thesis) to derive the concept that represents the system under study.

A key step in the ‘Good Modelling Practice’ (GMP) Handbook (Van Waveren et al., 1999) is step 3, “set up the model”, and more specifically step 3.4 on the construction of the conceptual model within the whole of the modelling process. This step helps to define the structure of the model and to explicitly formulate the reasons for the choices made in the modelling decision-making process. GMP asks for an explicit description of the general conceptual model, including (i) the domain of application (Chapters 3 and 5 of this thesis), (ii) the type of model required (i.e. explanatory vs. descriptive, Chapter 4), (iii) the relations and assumptions underlying the choices (Chapters 3, 4 and 5) and (iv) the verification of the conceptual model (Chapter 5).

Throughout the thesis, I emphasized the uncertainty associated with the model structure in compliance with the good modelling practices. It is important to identify and collect the key information on the different assumptions that led the modeller to select a particular model structure (Chapters 3 and 5). The protocol presented in Chapter 3 deals with these aspects, for instance through (i) using an uncertainty matrix (Walker et al., 2003; Refsgaard et al., 2006; 2007), (ii) clarifying the main shortcomings of the modelling solution, and (iii) revealing the main assumptions underlying the modelling approaches (i.e. modules, Belhouchette et al., 2009). Use of the protocol re-emphasizes the importance of considering the lack of certainty in our models (and knowledge) and helps to recognize the “unachievable” goal of a universal model (Van Oijen, 2009).

Dealing with this uncertainty starts with recognizing our incomplete knowledge. In this thesis, the incomplete knowledge is mostly identified through the use of the uncertainty matrix, but also via testing of different modelling approaches (e.g. Chapters 3 and 4), and discussing different hypotheses (Chamberlin, 1965). This methodology can be compared with the G2R3 approach (Grab-and-Glue, Run, Reject, Retry; Eldabi et al., 2005). It enables to quickly test a few hypotheses and draw conclusions from the outputs of the simulation exercise. It helps to reconstruct, repeat and reproduce the modelling process and to capitalize on the main outcomes of that process. It takes advantages of the “plug-and-play” facilities (Papajorgji, 2005) and enables to explicitly identify the validity domain of each of the modelling approaches, i.e. modules (Chapters 3 and 4), or modelling solutions when modules are coupled (Chapters 3 and 5). In the perspective of using a crop modelling framework, the process of model building should not be seen as a linear process (Rabbinge and De Wit, 1989), but rather as a cyclic process, that explicitly yields the uncertainty associated with each module tested.

We also identify the risk of compensating errors in the conceptual model (model structure) with errors in the parameter values. Such problems can be the reflection of misrepresentation of a component or of relations among components of the system. Indeed, modelling, by definition includes the process of simplifying a system, in which the relations between components are not fully known. Chapter 3 emphasizes the need to clearly identify the uncertainty in model structure related to our incomplete understanding. This was done through the use of an uncertainty matrix (Walker et al., 2003). Use of this matrix allows distinguishing different types and sources of uncertainties in order to facilitate uncertainty classification and to acknowledge that the underlying principles explaining the relations among components in the system

can never be fully known. Another example was given in Chapter 4 where the assumptions underlying the definition of the parameter representing the timing of the onset of leaf senescence, were different and not explicitly defined for each of the modelling approaches. Appropriate documentation of the modules and their underlying assumptions is essential (Donatelli et al., 2002; Bellochini et al., 2010; this thesis, e.g. Chapter 4).

Further, there is a need for an appropriate definition of the level of modelling detail in a process-based crop model to prevent “over-tuning” of parameters (Van Oijen, 2002; Chapter 4). Identification of the required level of detail is based on a number of pre-requisites, relying mostly on appropriate documentation of the decision-making process in model development (this thesis). Already Leffelaar (1990) pointed to the existence of an “optimum” level of detail in terms of the number of processes modelled that allows the closest approximation to system reality. In Chapter 4, we reinforced this idea and demonstrated that through an integrated use of complex (i.e. detailed) and simple (i.e. summarized) approaches more insight can be gained in how to model crop growth for large-scale applications, in support of decisions on the right trade-off between data availability and model detail. The omission of important relationships has the consequence of oversimplification in quantifying the radiation use efficiency (*RUE*). *RUE* is an aggregated value, measured over, usually, a period of, say, 10 days (or a growing cycle in Chapter 4), and it should therefore not be corrected for radiation and temperature on a daily basis (as done in models such as CropSyst, Stöckle et al., 2003), but rather on a seasonal basis.

Finally, we acknowledged the importance of visual tools as a way to communicate and clearly identify the validity domain corresponding to the specificity of the system. We discussed also the recent developments of declarative modelling as a good way forward (Chapter 5). With visual modelling environments such as SIMILE (Muetzelfeldt and Massheder, 2003), the modeller should be able to focus further on the modelling aspect rather than on the implementation/technical aspects. In Chapter 5, use of a conceptual model enabled to define what to include and what not to include in the model, by involving an expert to identify the basic crop processes that must be added or removed according to crop type.

Specific expertise is still needed to select the right modules for a specific simulation objective. However, in this thesis, by using crop modelling frameworks to go from modules to a MS, I stressed the importance of the documentation of the modelling process by (i) using conceptual modelling (e.g. visual tools) to define boundaries and

components to include in the model, (ii) emphasizing the role of the work group collaboration among crop experts and software engineers to create salient, legitimate and credible knowledge, and (iii) explicitly formulating the uncertainty related to the derived MS. These uncertainty assessments should be seen as the basis of the concept that enables to go from the objective of the simulation to the “right” modelling solution (Figure 6.2).

3. What will the future bring to crop modelling frameworks?

3.1 Requirements for successful development and use of crop modelling frameworks

3.1.1. The importance of clear modelling objectives

As Uran and Janssen (2003, p.525) mention, in general terms “All systems do what they were meant to do: they meet the functional specifications as defined at the start of the project. These specifications are usually rather vague and not clearly linked to the decision problem the system is designed to support”. Indeed, while developing a new tool (e.g. APES), the different partners involved may have different objectives in mind, although they may have agreed on the functional and technical requirements needed.

The development of a crop modelling framework (as defined in this thesis) is mostly targeted at supporting basic research aimed at increasing understanding of the functioning of the system, where simulation is used to synthesize the interactive effects of crop physiology, soil characteristics, crop management, and weather on system behaviour, testing different hypotheses. APES (in the framework of the SEAMLESS¹² project) was designed as a strategic tool within a system for policy assessment, to evaluate strategies and consequences of alternative crop systems, assessing the impacts of weather and management on production, water use, nutrient use, nutrient leaching, and economics. Some members of the APES development team however, considered APES more as a modular tool, for real-time operational decision support to assist in management decisions (sowing date, irrigation, fertilization, harvest date, pest management), putting emphasis on its modularity aspect, rather than on the integration of the tool within the broader SEAMLESS framework. As Van Delden et al. (2010) rightly state “developers are often focused on explaining the contents of the system, more than on its use value, while the interest of the users is often on the latter”. This mismatch of interests and objectives often results in a tool that can be

¹² SEAMLESS: System for Environmental and Agricultural Modelling: Linking European Science and Society; European project aiming at the development of a component-based framework in agricultural systems.

used by the developers themselves, but that is difficult to transfer to the intended users. Indeed, this statement is applicable to APES. All in all, APES fulfilled the functional requirements defined within the SEAMLESS project. However, despite a few workshops in which developers and users participated, currently only the developers can easily work with it, while other users still have difficulties.

3.1.2. Role of the crop modeller within an integrated project

From the beginning of crop model development, modellers had to combine knowledge of crop physiology, agronomy, modelling and programming skills (De Wit, 1968; Odum, 1983). However, with the recent advances in each of these disciplines, a gap has developed between the software engineers and the system analysts/crop physiologists/agronomists. For effective model development and application, there is an urgent need to reconcile these disciplines. In Chapter 5 an approach is described aiming at realizing this goal through the use of conceptual modelling to make the underlying assumptions in the model and the correspondence of concepts among the disciplines (basic crop processes = set of equations = modules) explicit. The modeller should play an essential role in bridging the gap between developers and users.

Developers¹³ of frameworks are responsible for the implementation of the modules and modelling solutions, they manage the implementation beam of the framework (architecture, design), and remain up to date with the ever faster advances in software engineer techniques (maintenance, evolution). End-users may use some modelling solutions for specific objectives in interaction with the modellers. Modellers are responsible for the actual development of the conceptual model to include in the framework (modules and modelling solutions) and should organise training courses, support and animations for the users, with assistance from the developers.

Consequently, the modeller should act as an interface between the developers and the end users: he/she must understand a minimum of all disciplines involved in the process of model development, integrating knowledge from the users (i.e. agronomists, crop physiologists) and the developers (software engineers) to bridge the gap (Jakeman et al., 2006; Chapter 5).

In the development of APES, the “work package” in charge of the biophysical modelling in the SEAMLESS project, the important role of the developer in creating the implementation beam of the framework came to the fore. However, very quickly the developers were confronted with the problem of testing and applying the modelling solutions created. Even if all the different modules incorporated in the

¹³ Developer, here, refers to software engineer.

modelling solution have been calibrated, the whole modelling solution will not necessarily yield “accurate” results. Consequently, within the SEAMLESS project, in response to these shortcomings, the ATILA group (Apes Testing and Improving Led by Applications, Casellas and Wery, 2009) was established.

The main objective of this group was to test and improve specific modelling solutions (i.e. combinations of modules and components) of the APES framework. The group included three types of members, corresponding to the three roles defined earlier: users with knowledge of agronomic and crop management aspects, i.e. soils, crops and agro-management and able to provide data for calibration and evaluation; modellers with knowledge of the components (model and variables), their expected behaviour and their interactions with other components; and the developers, capable of changing equations and parameters in the code version. Interaction among these three groups of experts helped in making progress and in creating a community spirit around the development of APES. APSIM is another good example, where a stable funding organization (APSRU, Agricultural Production Systems Research Unit in Australia) is supporting the development of the framework and the organization of regular training sessions for the users (<http://www.apsim.info>). The clear identification of these three roles and the identification of the appropriate experts responsible for each of these roles seem to be prerequisites for the success in the development of such types of modular tools.

3.2. Roles for crop modelling frameworks

3.2.1. Gaining insights in crop physiology by testing new hypotheses

It appears that conflicts between operationality and generality will always remain, and will affect the development and use of a crop modelling framework. Although the framework presented in this thesis rests on two main beams (the implementation beam and the modelling beam), the use of three different frameworks (i.e. implementation beams) in this thesis is a good example of the gap between what one could do (i.e. operationality) and what one wants to do (i.e. generality). Indeed, involvement of different research teams, different interests, and different objectives have resulted in the use of three implementation beams (i.e. software tools). However, it is my firm opinion that crop modelling frameworks (such as CROSPAL, APES, GCROP) represent promising research tools to deal with flexibility in crop modelling. The principles identified for the development of the modelling beam of the framework (Chapters 3, 4 and 5) remain applicable irrespective of the implementation beam used. These principles enable implementation of different approaches and support the

integration of current scientific knowledge (combining models in one tool) and the comparison of approaches to deepen our understanding of crop physiology.

A few examples of the application of crop modelling frameworks have been reported recently in the literature, especially using APSIM. For instance, Van Oosterom et al. (2006) substituted the module describing floral initiation with a gene network module. This type of work is essential for further development of a crop modelling framework to facilitate modification in the structure of crop models. A more recent example of the use of the modularity of APSIM is given by McMaster and Hargreaves (2009). They demonstrated the use of the principle of object-oriented design to couple principles of 3D modelling (using the phytomer concept as a unit repeated within and among shoots) with the plant module of APSIM. Following the same reasoning, it can be said that the use of a crop modelling framework can help in (i) creating and supporting a dialogue with crop physiologists and software engineers (Chapter 5) and (ii) identifying the impact of major process characteristics on crop growth factors, even at global scale (Chapter 4). These examples demonstrate that the use of crop modelling frameworks can accelerate the advancements in crop physiology in the modelling process, building on existing concepts.

3.2.2. Clarifying uncertainties on model outputs for integrated assessment

Crop modelling frameworks are increasingly used in integrated assessment studies, as originally APES was designed for in the framework of the SEAMLESS project. In such studies, the outputs of the crop model, assembled within the framework, are the interesting parts in the model chain, rather than the modelling process (Donatelli et al., 2010). Although integrated assessment studies also build on the advantages of modularity to create a model chain, the farm models only need indicators derived from the crop models as input variables (Janssen and Van Ittersum, 2007). Consequently, it is likely that when integrating a crop modelling framework (i.e. APES) in a larger modelling chain (e.g. SEAMLESS-Integrated Framework), the outputs of the models will be used by researchers not familiar with the biophysical models and their main underlying assumptions (Jakeman et al., 2006). Use of the modelling beam of the framework, as illustrated in this thesis, can help in reducing these drawbacks by increasingly involving the end user (i.e. farm modeller) in the model development process. Indeed, as the main focus of this thesis shows, the crop modelling framework appears a useful tool to deal with uncertainty in model structure.

Use of such a framework can be of great help in identifying the origins of the uncertainties and in interpreting the outcomes of the modelling solution that are used as inputs in other models in an integrated framework. As I am writing these lines, a new crop modelling framework¹⁴ is under development, building on the approaches used in the modular APES framework. The envisioned use of this new framework is mostly in integrated studies, coupling it with farm models (pers. comm. Van Ittersum, Group Plant Production Systems, Wageningen University, 2010). I trust that the concepts presented in this thesis will be taken into account and that the role of the crop modeller in the development and use of this new framework will be brought to the fore. Conceptual modelling and the uncertainty related to a specific modelling solution used should be prioritized, rather than software engineering techniques.

4. Conclusions: main contributions of this thesis in the crop modelling field

The development and use of a crop modelling framework has greatly contributed to the definition of guidelines that facilitate exchange of models (or parts of models, i.e. modules), representing different crop and cropping system processes, in dependence of user demands and objectives. The use of a crop modelling framework enables to (i) capitalize on new knowledge by testing alternative hypotheses without re-inventing the wheel, (ii) integrate different disciplines, and last but not least, (iii) communicate efficiently with the user of the tool by explicitly identifying the main uncertainties associated with its application.

Throughout the thesis, I have stressed the need for documentation of the modelling decision process to facilitate model reuse. Three main approaches have been used to explicitly document the process of decision-making for model building. The use of the uncertainty matrix emphasized the importance of explicitly defining the unknown. The use of sensitivity analysis enables tackling the issue of the required level of detail and highlights the risks of over-simplification of processes when data are scarce. The integration of expert knowledge in the development of the framework emphasizes the importance of explicitly describing the underlying assumptions through the use of conceptual modelling and the future potential of visual tools such as declarative modelling.

¹⁴ ACE: Analyze Cropping system and Environment.

To properly use crop modelling frameworks, there is a clear need to go back to the principle of conceptual modelling, either in its heuristic role (Hammer et al., 2002, Chapters 4 and 5) or for integrated studies (Van Delden et al., 2010, Chapter 3). While technical advances have stimulated substantial progress in the crop modelling field, especially in providing modular frameworks that allow easy coupling of different models at a higher scale for use in integrated assessment studies (Van Ittersum et al., 2008) or for further understanding of crop physiology (Hammer et al., 2002), conceptualisation of the systems remains an essential step. This thesis illustrates the continuing importance of the principles of systems analysis in the field of crop modelling, in combination with up-to-date advances in software engineering techniques: “crop modelling from conceptual modelling to software engineering and back”.



REFERENCES

- Acock, B., Reynolds, J.F., 1989. The rationale for adopting a modular generic structure for crop simulators. *Acta Horticulturae* 248, 391-400.
- Acock, B., Acock, M.C., 1991. Potential for using long-term field research data to develop and validate crop simulators. *Agronomy Journal* 83, 56-61.
- Acock, B., Reddy, V.R., 1997. Designing an object-oriented structure for crop models. *Ecological Modelling* 94, 33-44.
- Adam, M., Belhouchette, H., Casellas, E., Therond, O., Wery, J., 2009. APES, an agricultural production and externalities simulator evaluated for two main crops in Midi-Pyrenees. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), *AgSAP conference 2009*. Egmond aan Zee, The Netherlands, pp. 194-195.
- Adam M., Ewert F., Leffelaar P.A., Corbeels M., Van Keulen H., Wery J., 2010. CROSPAL, software that uses agronomic expert knowledge to assist modules selection for crop growth simulation. *Environmental Modelling & Software* 25, 946-955.
- Addiscott, T.M., 1998. Modelling concepts and their relation to the scale of the problem. *Nutrient Cycling in Agroecosystems* 50, 239-245.
- Ahuja L.R. M.L., and Howell T.A., 2002. Whole system integration and modeling- essential to agricultural Science and technology in the 21st century. In : Lajpat R. Ahuja L.R. M.L , Liwang Ma , and Terry A. Howell (Eds.), *Agricultural system models in the field research and technology transfer*, Lewis Publishers, Boca Raton, Florida, USA. pp. 1-7.
- Araus, J.L., Tapia, L., 1987. Photosynthetic Gas Exchange Characteristics of Wheat Flag Leaf Blades and Sheaths during Grain Filling: The Case of a Spring Crop Grown under Mediterranean Climate Conditions. *Plant Physiology* 85, 667-673.
- Argent, R.M., 2004. An overview of model integration for environmental applications-components, frameworks and semantics. *Environmental Modelling & Software* 19, 219-234.
- Belhouchette, H., Adam, M., Casellas, E., Celette, F., Corbeels, M., Wery, J., 2009. Performances of two crop models in various conditions: The importance of underlying assumptions. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), *AgSAP conference 2009*. Egmond aan Zee, The Netherlands, pp. 196-197.
- Bellocchi, G., Rivington, M., Donatelli, M., Matthews, K., 2010. Validation of biophysical models: issues and methodologies. A review. *Agronomy and Sustainable Development* 30, 109-130.
- Bender, J., Hertstein, U., Black, C.R., 1999. Growth and yield responses of spring wheat to increasing carbon dioxide, ozone and physiological stresses: a statistical analysis of 'ESPACE-wheat' results. *European Journal of Agronomy* 10, 185-195.
- Bergez, J.E., Colbach, N., Crespo, O., Garcia, F., Jeuffroy, M.H., Justes, E., Loyce, C., Munier-Jolain, N., Sadok, W., 2010. Designing crop management systems by simulation. *European Journal of Agronomy* 32, 3-9.
- Bindi, M., Porter, J.R., Miglietta, F., 1995. Comparison of Models to Simulate Leaf Appearance in Wheat. *European Journal of Agronomy* 4, 15-25.
- Bondeau, A., Smith, P.C., Zaehle, S., Schaphoff, S., Lucht, W., Cramer, W., Gerten, D., Lotze-Campen, H., Muller, C., Reichstein, M., Smith, B., 2007. Modelling the role of agriculture for the 20th century global terrestrial carbon balance. *Global Change Biology* 13, 679-706.
- Boogaard, H.L., Van Diepen, C.A., Rötter, R.P., Cabrera, J.M.C.A., Van Laar, H.H., 1998. WOFOST 7.1. User's guide for the WOFOST 7.1 crop growth simulation model and WOFOST Control Center 1.5. Technical Document 52, SC-DLO, Wageningen, The Netherlands. 144 pp.
- Boons-Prins, E.R., De Koning, G.H.J., Van Diepen, C.A., Penning de Vries, F.W.T., 1993. Crop-specific simulation parameters for yield forecasting across the European Community. Report CABO-DLO, Wageningen.
- Boote, K.J., Jones, J.W. , Nigel, P., 1996. Potential Uses and Limitations of Crop Models. *Agronomy Journal* 88, 704-716.

- Bouman, B.A.M., Van Keulen, H., Van Laar, H.H., Rabbinge, R., 1996. The 'School of de Wit' crop growth simulation models: A pedigree and historical overview. *Agricultural Systems* 52, 171-198.
- Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., 2003. An overview of the crop model STICS. *European Journal of Agronomy* 18, 309-332.
- Brooks, R.J., Tobias, A.M., 1996. Choosing the best model: Level of detail, complexity, and model performance. *Mathematical and Computer Modelling* 24, 1-14.
- Brun R., Kühni M., Siegrist H., Gujer W., Reichert P. 2002. Practical identifiability of ASM2d parameters--systematic selection and tuning of parameter subsets. *Water Resources* 36, 4113-4127.
- Casellas, E., Wery, J., 2009. Apes Testing and Improving Led by Applications (ATILA), Extra deliverable (WP6 taskforce, ATILA), 2009. SEAMLESS integrated project, EU 6th Framework programme, contract no. 010036-2, www.SEAMLESS-IP.org, 37 pp.
- Challinor, A.J., Wheeler, T.R., Craufurd, P.Q., Slingo, J.M., Grimes, D.I.F., 2004. Design and optimisation of a large-area process-based model for annual crops. *Agricultural and Forest Meteorology* 124, 99-120.
- Chamberlin, T.C., 1965. The Method of Multiple Working Hypotheses. *Science* 148, 754-759.
- Choudhury, B.J., 2000. A sensitivity analysis of the radiation use efficiency for gross photosynthesis and net carbon accumulation by wheat. *Agricultural and Forest Meteorology* 101, 217-234.
- Choudhury, B.J., 2001. Modeling radiation- and carbon-use efficiencies of maize, sorghum, and rice. *Agricultural and Forest Meteorology* 106, 317-330.
- Clements, P., 1995. From Subroutines to Subsystems: Component-Based Software Development. American Programmer, IEEE Computer Society 8(1),1-8
- Collatz, G.J., Ball, J.T., Grivet, C., Berry, J.A., 1991. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer. *Agricultural and Forest Meteorology* 54, 107-136.
- Collatz, G., Ribas-Carbo, M., Berry, J., 1992. Coupled Photosynthesis-Stomatal Conductance Model for Leaves of C4 Plants. *Australian Journal of Plant Physiology* 19, 519-538.
- Corbeels M., McMurtrie R.E., Pepper D.A., O'Connell A.M., 2005. A process-based model of nitrogen cycling in forest plantations: Part I. Structure, calibration and analysis of the decomposition model. *Ecological Modelling* 187, 426-448.
- Day, W., 2001. Modelling crop physiology for integrated decision making. *Annals of Applied Biology* 138, 215-219.
- De Willigen P., 1991. Nitrogen turnover in the soil-crop system; comparison of fourteen simulation models. *Nutrient Cycling in Agroecosystems* 27, 141-149.
- De Wit, C.T., 1965. Photosynthesis of leaf canopies. Verslagen Landbouwkundig Onderzoek (Agricultural Research Reports) nr. 663. Pudoc, Wageningen, The Netherlands.
- De Wit, C.T., 1968. Theorie en model. Veenman, Wageningen, The Netherlands.
- De Wit C.T. et al., 1978. Simulation of assimilation, respiration and transpiration of crops. *Simulation Monographs*, Pudoc, Wageningen, The Netherlands.
- De Wit C.T., 1992. Resource use efficiency in agriculture. *Agricultural Systems* 40, 125-151.
- De Wit, C.T., Brouwer, R., Penning De Vries, F.W.T., 1970. The simulation of photosynthetic systems. In: Setlik I., Editor, *Proceedings International Biological Program/Plant Production Technical Meeting Trebon, Prediction and management of photosynthetic productivity*, PUDOC, Wageningen, The Netherlands (1970), pp. 47-70.
- Debaeke, P., Nolot, J.-M., Raffaillac, D., 2006. A rule-based method for the development of crop management systems applied to grain sorghum in south-western France. *Agricultural Systems* 90, 180-201.

- Deschamps, M., Wery, J., 1987. Effet de l'eau et de l'azote sur la croissance et la nutrition azotée du pois et du pois chiche. In: Actes du colloque, Nutrition azotée des légumineuses, 19-21 novembre 1985, Versailles, France. INRA, Paris. pp. 270-280.
- Donatelli M., Van Ittersum M.K., Bindi M., Porter J.R., 2002. Modelling cropping systems-highlights of the symposium and preface to the special issues. *European Journal of Agronomy* 18:1-11.
- Donatelli, M., Carlini, L., Bellocchi, G., 2006a. A software component for estimating solar radiation. *Environmental Modelling & Software* 21, 411-416.
- Donatelli, M., Bellocchi, G., Carlini, L., 2006b. Sharing knowledge via software components: Models on reference evapotranspiration. *European Journal of Agronomy* 24, 186-192.
- Donatelli, M., Russell, G., Rizzoli, A.E., Acutis, M., Adam, M., Athanasiadis, I.N., Balderacchi, M., Luca Bechini, L., Belhouchette, H., Bellocchi, G., Bergez, J.E., Botta, M., Braudeau, E., Bregaglio, S., Carlini, L., Casellas, E., Celette, F., Ceotto, E., Charron-Moirez, M.H., Confalonieri, R., Corbeels, M., Criscuolo, L., Cruz, P., di Guardo, A., Ditto, D., Dupraz, C., Duru, M., Fiorani, D., Gentile, A., Ewert, F., Gary, C., Habyarimana, E., Jouany, C., Kansou, K., Knapen, R., Lanza Filippi, G., Leffelaar, P.A., Manici, L., Martin, G., Martin, P., Meuter, E., Mugueta, N., Mulia, R., Van Noordwijk, M., Oomen, R., Rosenmund, A., Rossi, V., Salinari, F., Serrano, A., Sorce, A., Vincent, G., Theau, J.P., Théron, O., Trevisan, M., Trevisiol, P., Van Evert, F.K., Wallach, D., Wery, J., Zerourou, A., 2010. A Component-Based Framework for Simulating Agricultural Production and Externalities. In: Brouwer, F., Van Ittersum, M.K. (Eds.), *Environmental and agricultural modelling: integrated approaches for policy impact assessment*. Springer, Dordrecht, The Netherlands. pp. 63-108.
- E.C., (Eds), 2001. A sustainable Europe for a better world: a European strategy for sustainable development. European Commission's proposal for the Gothenburg European Council, COM 264.
- Eldabi, T., Lee, M.W., Paul, R.J., 2005. An Evaluation of the World-Views for the Grab-and-Glue Framework. *Journal of Computing and Information Technology* 13 (2), 85-93.
- Ewert, F., 2004a. Modelling changes in global regionalized food production systems under changing climate and consequences for food security and environment - Development of an approach for improved crop modelling within IMAGE. Plant Production Systems Group, Department of Plant Sciences, Wageningen University & Netherlands Environmental Assessment Agency (MNP), National Institute for Public Health and Environment (RIVM), Wageningen/Bilthoven.
- Ewert, F., 2004b. Modelling Plant Responses to Elevated CO₂: How Important is Leaf Area Index? *Annals of Botany* 93, 619-627.
- Ewert, F., 2008. The (increasing) complexity of plant systems research and the use of models. In: Van Keulen, H., Van Laar, H.H., Rabbinge, R. (Eds.), *40 Years theory and model at Wageningen UR*. Wageningen University and Research Centre, Wageningen, The Netherlands. pp. 21-25.
- Ewert, F., Pleijel, H., 1999. Phenological development, leaf emergence, tillering and leaf area index, and duration of spring wheat across Europe in response to CO₂ and ozone. *European Journal of Agronomy* 10, 171-184.
- Ewert, F., Van Oijen, M., Porter, J.R., 1999. Simulation of growth and development processes of spring wheat in response to CO₂ and ozone for different sites and years in Europe using mechanistic crop simulation models. *European Journal of Agronomy* 10, 231-247.
- Ewert F., Rounsevell M.D.A., Reginster I., Metzger M.J., Leemans R., 2005. Future scenarios of European agricultural land use: I. Estimating changes in crop productivity. *Agriculture, Ecosystems & Environment* 107,101-116.

- Ewert, F., Van Ittersum, M.K., Bezlepina, I., Therond, O., Andersen, E., Belhouchette, H., Bockstaller, C., Brouwer, F., Heckelet, T., Janssen, S., Knapen, R., Kuiper, M., Louhichi, K., Olsson, J.A., Turpin, N., Wery, J., Wien, J.E., Wolf, J., 2009. A methodology for enhanced flexibility of integrated assessment in agriculture. *Environmental Science & Policy* 12, 546-561.
- Farquhar, G.D., Caemmerer, S., Berry, J.A., 1980. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species. *Planta* 149, 78-90.
- Feddes, R.A., Kowalik, P.J., Zaradny, H., 1978. Simulation of field water use and crop yield, *Simulation Monographs*, Pudoc, Wageningen, The Netherlands. 189 pp.
- Forrester, J.W., 1961. *Industrial dynamics*. MIT Press, Boston, USA.
- Fowler M., 1997. *Analysis patterns : reusable object models*. Addison-Wesley, Menlo Park, California, USA.
- Gamma, E., Helm, R., Johnson, R., Vlissides, J., 1995. *Design Patterns: elements of reusable object-oriented software*. Addison-Wesley, Boston, Massachusetts, USA.
- Giunta, F., Pruneddu, G., Motzo, R., 2009. Radiation interception and biomass and nitrogen accumulation in different cereal and grain legume species. *Field Crops Research* 110, 76-84.
- Goudriaan, J., Van Laar H.H., 1994. *Modelling potential crop growth processes: textbook with exercises*. Kluwer, Dordrecht.
- Green C.F., 1987. Nitrogen nutrition and wheat growth in relation to absorbed solar radiation. *Agricultural & Forest Meteorology* 41, 207-248.
- Greenwood D.J., Wood J.T., Cleaver T.J., 1974. A dynamic model for the effects of soil and weather conditions on nitrogen response. *The Journal of Agricultural Science* 82, 455-467.
- Gregory, P.J., 1998. Alternative crops for duplex soils: Growth and water use of some cereal, legume, and oilseed crops, and pastures. *Australian Journal of Agricultural Research* 49, 21-32.
- Groot J.J.R., Verberne E.L.J., 1991. Response of wheat to nitrogen fertilization, a data set to validate simulation models for nitrogen dynamics in crop and soil. *Nutrient Cycling in Agroecosystems* 27, 349-383.
- Guilioni, L., Wery, J., Lecoecur, J., 2003. High temperature and water deficit may reduce seed number in field pea purely by decreasing plant growth rate. *Functional Plant Biology* 30, 1151-1164.
- Hammer, G.L., 1998. Crop modelling: current status and opportunities to advance. *Acta Horticulturae* 456, 27-36.
- Hammer, G.L., Stephens, D.J., Butler, D.J., 1996. Development of a national drought alert strategic information system - Wheat modelling sub-project (a) Development of predictive models of wheat production. Final report to LWRRDC, 41 pp.
- Hammer, G.L., Kropff, M.J., Sinclair, T.R., Porter, J.R., 2002. Future contributions of crop modelling--from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. *European Journal of Agronomy* 18, 15-31.
- Hansen, J.W., Jones, J.W., 2000. Scaling-up crop models for climate variability applications. *Agricultural Systems* 65, 43-72.
- Havelka, U.D., Wittenback, V.A., Boyle, M.G., 1984. CO₂-Enrichment Effects on Wheat Yield and Physiology. *Crop Science* 24, 1163-1168.
- Haxeltine, A., Prentice, I.C., 1996a. BIOME3: An Equilibrium Terrestrial Biosphere Model Based on Ecophysiological Constraints, Resource Availability, and Competition Among Plant Functional Types. *Global Biogeochemical Cycles* 10, 693-709.
- Haxeltine, A., Prentice, I.C., 1996b. A General Model for the Light-Use Efficiency of Primary Production. *Functional Ecology* 10, 551-561.
- Hay, R.K.M., Porter, J.R., 2006. *The physiology of crop yield*. Blackwell Publishing, Second Edition. Blackwell Publishing, Ames, Iowa, USA. 328 pp.

- Hearn, A.B., 1994. OZCOT: A simulation model for cotton crop management. *Agricultural Systems* 44, 257-299.
- Heath, O.V.S., Gregory, F.G., 1938. The constancy of the mean net assimilation rate and its ecological importance. *Annals of Botany* 2, 811-818.
- Heemskerk, M., Wilson, K., Pavao-Zuckerman, M., 2003. Conceptual models as tools for communication across disciplines. *Conservation Ecology* 7(3), 8-21
- Hillyer, C., Bolte, J., Van Evert, F., Lamaker, A., 2003. The ModCom modular simulation system. *European Journal of Agronomy* 18, 333-343.
- Holzworth, D., Huth, N., 2009. Reflection + XML simplifies development of the APSIM generic PLANT model. In: Anderssen, R.S., Braddock, R.D., Newham, L.T.H. (Eds.), 18th World IMACS congress and MODSIM09 international congress on modelling and simulation. Modelling and simulation society of Australia and New Zealand and international association for mathematics and computers in simulation, July 2009, pp. 887-893.
- Holzworth, D.P., Huth, N.I., de Voil, P.G., 2010. Simplifying environmental model reuse. *Environmental Modelling & Software* 25, 269-275.
- Hooda, R.S., Sheoran, I.S., Singh, R., 1989. Ontogenetic changes in photosynthesis, respiration, nitrogen-fixation and water-use efficiency in Chickpea (*Cicer arietinum* L.) grown at two moisture levels. *Photosynthetica* 23, 189-196.
- Horsch E., 2000. The effect of nitrogen and plant density on canopy development, crop growth and radiation use efficiency in maize. MSc. Thesis, Wageningen University, Wageningen, The Netherlands.
- Jagtap, S.S., Jones, J.W., 2002. Adaptation and evaluation of the CROPGRO-soybean model to predict regional yield and production. *Agriculture, Ecosystems and Environment* 93, 73-85.
- Jakeman, A.J., Letcher, R.A., Norton, J.P., 2006. Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software* 21, 602-614.
- Jamieson, P.D., Brooking, I.R., Porter, J.R., Wilson, D.R., 1995. Prediction of leaf appearance in wheat: a question of temperature. *Field Crops Research* 41, 35-44.
- Jamieson P.D., Semenov M.A., Brooking I.R., Francis G.S., 1998. Sirius: a mechanistic model of wheat response to environmental variation. *European Journal of Agronomy* 8, 161-179.
- Jamieson, P.D., Porter, J.R., Goudriaan, J., Ritchie, J.T., Van Keulen, H., Stol, W., 1998. A comparison of the models AFRCWHEAT2, CERES-Wheat, Sirius, SUCROS2 and SWHEAT with measurements from wheat grown under drought. *Field Crops Research* 55, 23-44.
- Janssen, S., Van Ittersum, M.K., 2007. Assessing farm innovations and responses to policies: A review of bio-economic farm models. *Agricultural Systems* 94, 622-636.
- Johnsson H., Bergstrom L., Jansson P.-E., Paustian K., 1987. Simulated nitrogen dynamics and losses in a layered agricultural soil. *Agriculture, Ecosystems & Environment* 18, 333-356.
- Jones J.W., Hoogenboom G., Porter C.H., Boote K.J., Batchelor W.D., Hunt L.A., Wilkens P.W., Singh U., Gijsman A.J., Ritchie J.T., 2003. The DSSAT cropping system model. *European Journal of Agronomy* 18, 235-265.
- Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. *European Journal of Agronomy* 18, 235-265.
- Kaschuk, G., 2009. Sink stimulation of leaf photosynthesis by the carbon costs of rhizobial and arbuscular mycorrhizal fungal symbioses. PhD Thesis. Wageningen University, Wageningen, The Netherlands.
- Kaschuk, G., Kuiper, T.W., Leffelaar, P.A., Hungria, M., Giller, K.E., 2009. Are the rates of photosynthesis stimulated by the carbon sink strength of rhizobial and arbuscular mycorrhizal symbioses? *Soil Biology and Biochemistry* 41, 1233-1244.

- Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., 2003. An overview of APSIM, a model designed for farming systems simulation. *European Journal of Agronomy* 18, 267-288.
- Kimball, B.A., Jr., P.J.P., Garcia, R.L., LaMorte, R.L., Wall, G.W., Hunsaker, D.J., Wechsung, G., Wechsung, F., Kartschall, T., 1995. Productivity and water use of wheat under free-air CO₂ enrichment. *Global Change Biology* 1, 429-442.
- Kimball, B.A., LaMorte, R.L., Pinter, P.J., Jr., Wall, G.W., Hunsaker, D.J., Adamsen, F.J., Leavitt, S.W., Thompson, T.L., Matthias, A.D., Brooks, T.J., 1999. Free-Air CO₂ enrichment and soil nitrogen effects on energy balance and evapotranspiration of wheat. *Water Resources Research* 35, 1179-1190.
- Kirby, E.J.M., Siddique, K.H.M., Perry, M.W., Kaesehagen, D., Stern, W.R., 1989. Variation in Spikelet Initiation and Ear Development of Old and Modern Australian Wheat Varieties. *Field Crops Research* 20, 113-128.
- Lecoeur J., Guilioni L., 1998. Rate of leaf production in response to soil water deficits in field pea. *Field Crops Research* 57, 319-328.
- Lecoeur J., Wery J., Turc O., 1992. Osmotic adjustment as a mechanism of dehydration postponement in chickpea (*Cicer arietinum* L.) leaves. *Plant and Soil* 144, 177 - 189.
- Leemans, R., 1997. Effects of Global Change on Agricultural Land Use: Scaling Up from Physiological Processes to Ecosystem Dynamics. *Ecology in Agriculture*. Academic Press, pp. 415-452.
- Leffelaar, P.A., 1990. On scale problems in modelling: an example from soil ecology. In: Rabbinge, R., Goudriaan, J., Van Keulen, H., Van Laar, H.H., Penning de Vries, F.W.Y. (Eds.), *Theoretical Production Ecology: reflections and prospects*. Simulation Monographs 34, Pudoc, Wageningen, The Netherlands. pp. 57-73.
- Leffelaar, P.A. (Ed.), 1999. On systems analysis and simulation of ecological processes; with examples in CSMP, FST and FORTRAN. Second Edition. *Current Issues in Production Ecology*, Vol 4. Kluwer Academic Publishers, Dordrecht, The Netherlands. 318 pp.
- Leffelaar P.A., Van Evert F., Corbeels M., Celette F., Casellas E., 2007. Is there a large difference between the nitrogen uptake by roots and by microbes? Personal communication in Prague, 19 April, 2007.
- Lemaire, G., 1997. *Diagnosis of the nitrogen status in crops*. Springer, Berlin, Germany. 241 pp.
- Liu, J., Williams, J.R., Zehnder, A.J.B., Yang, H., 2007. GEPIC - modelling wheat yield and crop water productivity with high resolution on a global scale. *Agricultural Systems* 94, 478-493.
- Louhichi, K., Janssen, S., Li, H., Borkowski, N., Kanellopoulos, A., Flichman, G., Hengsdijk, H., Zander, P., Blanco Fonesca, M., Athanasiadis, I.N., Rizzoli, A.E., Huber, D., Van Ittersum, M.K., 2010. A generic Farming System Simulator (FSSIM); an application for modelling European arable farming. In: Brouwer, F.M., Van Ittersum, M. K. (Eds.), *Environmental and agricultural modelling: integrated approaches for policy impact assessment*. Springer, Dordrecht, the Netherlands. pp. 109-133.
- Mahieu, S., Germon, F., Aveline, A., Hauggaard-Nielsen, H., Ambus, P., Jensen, E.S., 2009. The influence of water stress on biomass and N accumulation, N partitioning between above and below ground parts and on N rhizodeposition during reproductive growth of pea (*Pisum sativum* L.). *Soil Biology and Biochemistry* 41, 380-387.
- Mahmood, F., Mallouli, A., Belhouchette, H., Wery, J., 2010. Impact assessment of the introduction of grain legumes in cereal-based cropping systems in the Midi-Pyrénées region (France) using the modelling chain: APES-FSSIM-Indicators. *Proceedings of AGRO2010*, ESA (eds).
- Mayer, D.G., Butler, D.G., 1993. Statistical validation. *Ecological Modelling* 68, 21-32.
- McCown R.L., Hammer G.L., Hargreaves J.N.G., Holzworth D.P., Freebairn D.M., 1996. APSIM: a novel software system for model development, model testing and simulation in agricultural systems research. *Agricultural Systems* 50, 255-271.

- McMaster, G.S., Hargreaves, J.N.G., 2009. CANON in D(esign): Composing scales of plant canopies from phytomers to whole-plants using the composite design pattern. *NJAS - Wageningen Journal of Life Sciences* 57, 39-51.
- Medlyn, B.E., 1998. Physiological basis of the light use efficiency model. *Tree Physiology* 18, 167-176.
- Meinke, H., Hammer, G.L., Van Keulen, H., Rabbinge, R., Keating, B.A., 1997. Improving wheat simulation capabilities in Australia from a cropping systems perspective: Water and nitrogen effects on spring wheat in a semi-arid environment. *European Journal of Agronomy* 7, 75-88.
- Metzger, M.J., Bunce, R. G. H., Jongman, R.H.G., Mûcher, C.A., Watkins, J.W., 2005. A climatic stratification of the environment of Europe. *Global Ecology & Biogeography* 14, 549-563.
- Mi, G., Tang, L., Zhang, F., Zhang, J., 2000. Is nitrogen uptake after anthesis in wheat regulated by sink size? *Field Crops Research* 68, 183-190.
- Miglietta, F., 1992. Simulation of wheat ontogenesis. Ph.D.Thesis, Wageningen Agricultural University, Wageningen, The Netherlands.
- Miglietta, F., Porter, J.R., 1992. The Effects of Climatic Change on Development in Wheat: Analysis and Modelling. *Journal of Experimental Botany* 43, 1147-1157.
- Mitchell, R., Manderscheid, R., Fereres, E., Rodriquez, D., Ewert, F., Semenov, M., Burkart, S., Goudriaan, J., Schütz, M., Fangmeier, A., Jamieson, P., Porter, J., Richter, G., Ruz, C., Gómez, H., Testi, L., Villalobos, F., Weigel, H., 2000. IMPETUS Final report: Improving wheat model accuracy and suitability for regional impact assessment. EC Project ENV4-CT97-0496. 150 pp.
- Monteith, J.J.L., 1977. Climate and the efficiency of crop production in Britain. *Philosophical Transactions of the Royal Society of London B* 281, 277-294.
- Monteith, J.L., 1996. The Quest for Balance in Crop Modeling. *Agronomy Journal* 88, 695-697.
- Moore, A.D., Holzworth, D.P., Herrmann, N.I., Huth, N.I., Robertson, M.J., 2007. The Common Modelling Protocol: A hierarchical framework for simulation of agricultural and environmental systems. *Agricultural Systems* 95, 37-48.
- Muetzelfeldt, R., Massheder, J., 2003. The Simile visual modelling environment. *European Journal of Agronomy* 18, 345-358.
- Mulia, R., Dupraz, C., 2006. Unusual fine root distributions of two deciduous tree species in southern France: what consequences for modelling of tree root dynamics? *Plant Soil* 281, 71-85.
- Munier-Jolain, N., Biarnès, V., Chaillet, I., Lecoœur, J., Jeuffroy, M.H., 2005. Agrophysiology du pois protéagineux. *Physiology of the Protein Pea*. INRA, ARVALIS, Paris. 284 pp.
- Munns R., Pearson C., 1974. Effect of water deficit on translocation of carbohydrate in *Solanum tuberosum*. *Functional Plant Biology* 1, 529-537.
- Neitsch, S.L., Arnold J.G., Kiniry J.R., Srinivasan R., Williams J.R., 2002. Soil and Water Assessment Tool User's Manual, Version 2000. USDA-ARS-SR Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, Texas, US. 506 pp.
- Neitsch, S.L., Arnold, J.G., Kiniry, J.R., Williams, J.R., 2005. Soil and Water Assessment Tool Theoretical Documentation Version 2005, USDA-ARS-SR Grassland, Soil and Water Research Laboratory, Agricultural Research Service, Temple, Texas, US. 494 pp.
- Ney, B., Duthion, C., Turc, O., 1994. Phenological response of pea to water stress during reproductive development. *Crop Science* 34, 141-146.
- Odum, H.T., 1983. *Systems Ecology: An Introduction*. (Eds), Wiley, New York, USA. 644 pp.
- Olesen, J.E., Petersen, B.M., Berntsen, J., Hansen, S., Jamieson, P.D., Thomsen, A.G., 2002. Comparison of methods for simulating effects of nitrogen on green area index and dry matter growth in winter wheat. *Field Crops Research* 74, 131-149.
- Papajorgji, P., 2005. A plug and play approach for developing environmental models. *Environmental Modelling & Software* 20, 1353-1357.
- Papajorgji P.J., Pardalos P.M., 2006. Software engineering techniques applied to agricultural systems.

- Passioura, J.B., 1996. Simulation models: science, snake oil, education, or engineering? *Agronomy Journal* 88, 690-694.
- Penman, H.L., 1948. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society A* 193, 120-145.
- Penning de Vries, F.W.T., 1974. Substrate utilization and respiration in relation to growth and maintenance in higher plants *Netherlands Journal Agric. Sci.* 22, 40-44.
- Penning de Vries, F.W.T., 1982. Systems analysis and models of crop growth. In: Penning de Vries, F.W.T., Van Laar, H.H. (Eds.) *Simulation of plant growth and crop production*, Simulation Monographs, Pudoc, Wageningen. pp. 9-25.
- Pronk A., Goudriaan J., Stilma E., Challa H., 2003. A simple method to estimate light interception by nursery stock conifers: a case study of eastern white cedar. *Netherlands Journal of Agricultural Sciences* 51, 279-295
- Quesnel, G., Duboz, R., Ramat, É., 2009. The Virtual Laboratory Environment - An operational framework for multi-modelling, simulation and analysis of complex dynamical systems. *Simulation Modelling Practice and Theory* 17, 641-653.
- Qureshi, M.R.J., Hussain, S.A., 2008. A reusable software component-based development process model. *Advances in Engineering Software* 39, 88-94.
- Rabbinge, R., De Wit, C.T., 1989. Systems, models and simulation. In: Rabbinge, R., Ward, S.A., Van Laar, H.H. (Eds.), *Simulation and systems management in crop protection*. Simulation Monographs, Pudoc, Wageningen, The Netherlands. pp. 3-15.
- Rahman, J.M., Seaton, S.P., Cuddy, S.M., 2004. Making frameworks more useable: using model introspection and metadata to develop model processing tools. *Environmental Modelling & Software* 19, 275-284.
- Rappoldt, C., Van Kraalingen, D.W.G., 1996. The Fortran Simulation Translator. FST version 2.0. Introduction and Reference Manual. Quantitative Approaches in Systems Analysis No. 5, C.T. de Wit Graduate School for Production Ecology, Wageningen University, Wageningen, The Netherlands, 178 pp.
- Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V., Hobbie, J.E., 1992. Aggregating Fine-Scale Ecological Knowledge to Model Coarser-Scale Attributes of Ecosystems. *Ecological Applications* 2, 55-70.
- Refsgaard, J.C., Van der Sluijs, J.P., Brown, J., Van der Keur, P., 2006. A framework for dealing with uncertainty due to model structure error. *Advances in Water Resources* 29, 1586-1597.
- Refsgaard J.C., Van der Sluijs J.P., Højberg A.L., Vanrolleghem P.A., 2007. Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software* 22, 1543-1556.
- Reidsma, P., Ewert, F., Oude Lansink, A., Leemans, R., 2010. Adaptation to climate change and climate variability in European agriculture: The importance of farm level responses. *European Journal of Agronomy* 32, 91-102.
- Reynolds, J.F., Acock, B., 1997. Modularity and genericness in plant and ecosystem models. *Ecological Modelling* 94, 7-16.
- Ritchie, J.T. and Otter, S., 1985. Description and Performance of CERES-wheat: a User-oriented Wheat Yield Model. ARS Wheat Yield Project, USDA-ARS, Vol. 38, Chapter 10, pp. 159-175.
- Roberts, E.H., Summerfield, R.J., 1987. Measurement and prediction of flowering in annual crops. In: Atherton, J.G. (Ed.), *Manipulation of flowering*. Butterworths, London, UK, pp. 17-50.
- Roux, S., Adam, M., Wery, J., 2010. On an interdisciplinary framework for the development and analysis of crop models. In proceeding iEMSs, International Congress on Environmental Modelling and Software, Modelling for Environment's Sake, Ottawa, Canada.

- Russell, G., Wilson, G.W., 1994. An agro-pedo-climatological knowledge-base of wheat in Europe. Office for Official Publications of the European Communities, ECSC-EC-EAEC, Agric.Series CL-NA-15789-EN-C. Brussels, Luxembourg. 158 pp
- Salinari F., Rossi V., Manici L.M., 2008. A model framework for simulation plant disease epidemics, in: P. R. Pisa (EdS.), The Xth European Society of Agronomy Congress, Bologna, Italy. pp. 751-753.
- Scholten, H., 2008. Better modelling practice: an ontological perspective on multidisciplinary, model-based problem solving. Ph.D Thesis Wageningen University, Wageningen, 314 pp.
- Sharpley A.N., Williams J.R., 1990. EPIC-Erosion Productivity Impact Calculator. Model Documentation. U. S. Department of Agriculture Technical Bulletin 1.
- Shibu M.E., Leffelaar P.A., Van Keulen H., Aggarwal P.K., 2010. LINTUL3, a simulation model for nitrogen-limited situations: Application to rice. *European Journal of Agronomy* 32, 255-271.
- Sinclair T. Seligman N.G., 1996. Crop Modeling: From Infancy to Maturity. *Agronomy Journal* 88, 698-704.
- Sinclair T.R., Amir J., 1992. A model to assess nitrogen limitations on the growth and yield of spring wheat. *Field Crops Research* 30, 63-78.
- Sinclair T.R., Muchow R.C., 1999. Radiation Use Efficiency, *Advances in Agronomy*, Academic Press. pp. 215-265.
- Sitch, S., B. Smith, I.C. Prentice, A. Arneth, A. Bondeau, W. Cramer, J.O. Kaplan, S. Levis, W. Lucht, M.T. Sykes, K. Thonicke, S. Venevsky. 2003. Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology* 9, 161-185.
- Slafer, G.A., Rawson, H.M., 1994. Sensitivity of Wheat Phasic Development to Major Environmental-Factors - a Reexamination of Some Assumptions Made by Physiologists and Modelers. *Australian Journal of Plant Physiology* 21, 393-426.
- Spitters C.J.T., Schapendonk A.H.C.M., 1990. Evaluation of breeding strategies for drought
- Spitters C.J.T., 1987. An analysis of variation in yield among potato cultivars in terms of light absorption, light utilization and dry matter partitioning, *Acta Horticulturae* 214, 71-84.
- Spitters, C.J.T., 1990. Crop growth models: their usefulness and limitations. *Acta Horticulturae* 267, 349-368.
- Sterk, B., Leeuwis, C., Van Ittersum, M.K., 2009. Land use models in complex societal problem solving: Plug and play or networking? *Environmental Modelling & Software* 24, 165-172.
- Stöckle C.O., Martin S.A., Campbell G.S., 1994. CropSyst, a cropping systems simulation model: Water/nitrogen budgets and crop yield. *Agricultural Systems* 46, 335-359.
- Stöckle, C.O., Donatelli, M., Nelson, R., 2003. CropSyst, a cropping systems simulation model. *European Journal of Agronomy* 18, 289-307.
- Stöckle, C.O., Kemanian, A.R., 2009. Crop Radiation Capture and Use Efficiency: A Framework for Crop Growth Analysis, *Crop Physiology*. Academic Press, San Diego, pp. 145-170.
- Streck, N.A., Weiss, A., Xue, Q., Stephen Baenziger, P., 2003. Improving predictions of developmental stages in winter wheat: a modified Wang and Engel model. *Agricultural & Forest Meteorology* 115, 139-150.
- Supit, I., Hooijer, A.A., Van Diepen, C.A., (Eds), 1994. System description of the WOFOST 6.0 crop simulation model implemented in CGMS. Publication EUR 15956 EN of the Office for Official Publications of the EU, Luxembourg.
- Tan, G., Shibasaki, R., 2003. Global estimation of crop productivity and the impacts of global warming by GIS and EPIC integration. *Ecological Modelling* 168, 357-370.

- Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., Oomen, R., Russell, G., Ewert, F., Bergez, J.-E., Janssen, S., Wery, J., Van Ittersum, M.K., 2010. Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. *Agriculture, Ecosystems & Environment*, In Press, Corrected Proof, Available online 17 June 2010.
- Tittonell, P., 2008. Msimu wa Kupanda : Targeting resources within diverse, heterogenous and dynamic farming systems of East Africa. Ph.D. Thesis, Wageningen University, Wageningen. 320 pp.
- Turc, O., Wery, J., Cheong, G.S.C., 1990. Drought stress as a tool for improvement of pea seed production. In: Scaife, A. (Ed.), *European Society of Agronomy Inaugural Congress*, Paris.
- Uran, O., Janssen, R., 2003. Why are spatial decision support systems not used? Some experiences from the Netherlands. *Computers, Environment and Urban Systems* 27, 511-526.
- Van Daalen, C.E., Dresen, L., Janssen, M.A., 2002. The roles of computer models in the environmental policy life cycle. *Environmental Science & Policy* 5, 221-231.
- Van Delden A., Kropff M.J., Haverkort A.J., 2001. Modeling temperature- and radiation-driven leaf area expansion in the contrasting crops potato and wheat. *Field Crops Research* 72:119-141.
- Van Delden, H., Seppelt, R., White, R., Jakeman, A.J., 2010. A methodology for the design and development of integrated models for policy support. *Environmental Modelling & Software* In Press, Corrected Proof.
- Van der Schoot J.R., Neuvel J., Van Dijk W., 2002. Interactie stikstof- en fosfaatvoorziening bij aardappel, *Praktijkonderzoek Plant & Omgeving B.V.*, Sector agv.
- Van Evert, F.K., Holzworth, D., Muetzelfeldt, R.M., Rizzoli, A.E., Villa, F., 2005. Convergence in integrated modeling frameworks. In: *Proceedings of the MODSIM 2005 International Congress on Modelling and Simulation*. Modelling and Simulation Society of Australia/New Zealand, Melbourne, Australia.
- Van Ittersum, M.K. , Rabbinge, R., 1997. Concepts in production ecology for analysis and quantification of agricultural input-output combinations. *Field Crops Research* 52, 197-208.
- Van Ittersum, M.K., Leffelaar, P.A., Van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. *European Journal of Agronomy* 18, 201-234.
- Van Ittersum, M.K., Ewert, F., Heckeleei, T., Wery, J., Alkan Olsson, J., Andersen, E., Bezlepkina, I., Brouwer, F., Donatelli, M., Flichman, G., Olsson, L., Rizzoli, A.E., Van der Wal, T., Wien, J.E., Wolf, J., 2008. Integrated assessment of agricultural systems - A component-based framework for the European Union (SEAMLESS). *Agricultural Systems* 96, 150-165.
- Van Keulen, H., Penning de Vries, F.W.T., Drees, E.M., 1982. A summary model for crop growth. In: Penning de Vries, F.W.T., Van Laar, H.H. (Eds.), *Simulation of plant growth and crop production*. Simulation Monographs, Pudoc, Wageningen, pp. 87-97.
- Van Keulen, H., Wolf, J., (Eds.), 1986. *Modelling of agricultural production: weather, soils and crops*. Simulation Monographs, Pudoc, Wageningen, The Netherlands.
- Van Keulen, H., Seligman, N.G., 1987. Simulation of water use, nitrogen nutrition and growth of a spring wheat crop. *Simulation Monographs*, Pudoc, Wageningen, The Netherlands, 310 pp.
- Van Kraalingen, D.W.G., 1990. User manual of CABO/TPE Weather system. CABO/TPE internal communication, WAU-TPE, Wageningen.
- Van Kraalingen, D.W.G., 1995. The FSE-system for crop simulation, version 2.1. *Quantitative Approaches in Systems Analysis No. 1*, C.T. de Wit Graduate School for Production Ecology, Wageningen University, Wageningen, The Netherlands, 58 pp.
- Van Oijen M., 2002. On the use of specific publication criteria for papers on process-based modelling in plant science. *Field Crops Research* 74, 197-205.

- Van Oijen, M., 2009. Theory and models for managed ecosystems: from confusion to certainty and back again. In: Van Keulen, H., Van Laar, H.H., Rabbinge, R. (Eds.), 40 Years theory and model at Wageningen UR. Wageningen University and Research Centre, Wageningen, The Netherlands. pp. 25-32.
- Van Oijen, M., Schapendonk, A.H.C.M., Jansen, M.J.H., Pot, C.S., Van Kleef, J., Goudriaan, J., 1998. Effects of elevated CO₂ on development and morphology of spring wheat grown in cooled and non-cooled open-top chambers. *Australian Journal of Plant Physiology* 25, 617-626.
- Van Oijen, M., Ewert, F., 1999. The effects of climatic variation in Europe on the yield response of spring wheat cv. Minaret to elevated CO₂ and O₃: an analysis of open-top chamber experiments by means of two crop growth simulation models. *European Journal of Agronomy* 10, 249-264.
- Van Oosterom, E., Hammer, G., Chapman, S., Doherty, A., 2006. A simple gene network model for photoperiodic response of floral transition in sorghum can generate genotype-by-environment interactions in grain yield at the crop level. In: C.F. Mercer (Ed.), *Breeding for Success: Diversity in Action*. Proceedings of the 13th Australasian Plant Breeding Conference, Christchurch, New Zealand, 18-21 April 2006. pp. 687-691. CDROM format (ISBN 978-0-86476-167-8).
- Van Waveren, R.H., Groot S., Scholten H., Van Geer F., Wösten H., Koeze R., Noort J., 1999. *Good Modelling Practice Handbook*, STOWA, Utrecht, RWS-RIZA, Lelystad, The Netherlands (in Dutch, English version from <http://www.landandwater.tudelft.nl/downloads/GMP-uk.pdf>).
- Villa, F., 2007. A semantic framework and software design to enable the transparent integration, reorganization and discovery of natural systems knowledge. *Journal of Intelligent Information Systems* 29, 79-96.
- Voinov, A., Fitz, C., Boumans, R., Costanza, R., 2004. Modular ecosystem modeling. *Environmental Modelling & Software* 19, 285-304.
- Voisin, A.-S., Bourion, V., Duc, G., Salon, C., 2007. Using an Ecophysiological Analysis to Dissect Genetic Variability and to Propose an Ideotype for Nitrogen Nutrition in Pea. *Annals of Botany* 100, 1525-1536.
- Vos J., Putten P.E.L.v.d., Birch C.J., 2005. Effect of nitrogen supply on leaf appearance, leaf growth, leaf nitrogen economy and photosynthetic capacity in maize (*Zea mays* L.). *Field Crops Research* 93, 64-73.
- Walker W.E., Harremoes P., Rotmans J., Van der Sluijs J.P., Van Asselt M.B.A., Janssen P., Kraymer von Krauss M.P., 2003. Defining uncertainty: a conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment* 4, 5-17.
- Wallach, D.E., Makowski, D., Jones, J.W. (Eds.), 2006. *Working with dynamic crop models: evaluating, analyzing, parameterizing, and applications*. Elsevier, Amsterdam, The Netherlands, 447 pp.
- Wang, E., Engel, T., 2000. SPASS: a generic process-oriented crop model with versatile windows interfaces. *Environmental Modelling & Software* 15, 179-188.
- Wang, E., Robertson, M.J., Hammer, G.L., Carberry, P.S., Holzworth, D., Meinke, H., Chapman, S.C., Hargreaves, J.N.G., Huth, N.I., McLean, G., 2002. Development of a generic crop model template in the cropping system model APSIM. *European Journal of Agronomy* 18, 121-140.
- Watson, D.J., 1947. Comparative Physiological Studies on the Growth of Field Crops: I. Variation in Net Assimilation Rate and Leaf Area between Species and Varieties, and within and between Years. *Annals of Botany* 11, 41-76.
- Wery, J., 1996. Production de graines et fixation d'azote par des cultures de légumineuses sous contrainte hydrique. Habilitation à diriger des recherches, Montpellier, France, formation doctoral: bases de la production végétale, Université Montpellier II Sciences et techniques du Languedoc. 65 pp.

- Wery, J., 2005. Differential effects of soil water deficit on the basic plant functions and their significance to analyse crop responses to water deficit in indeterminate plants. *Australian Journal of Agricultural Research* 56, 1201-1209.
- Wery, J., Turc, O., Salsac, L., 1986. Relationship between growth, nitrogen fixation and assimilation in a legume (*Medicago sativa* L.). *Plant Soil* 96, 17-29.
- Wery, J., Deschamps, M., Leger-Cresson, N., 1988. Influence of some agroclimatic factors and agronomic practices on nitrogen nutrition of chickpea (*Cicer arietinum* L.). In: *Nitrogen fixation by legumes in Mediterranean agriculture* (Beck, D.P., Materon, L.A., Eds.). ICARDA, Aleppo, Syria. pp. 287-301.
- Wery, J., Ahlawat, I.P.S., 2008. Analysing and improving the role of grain legumes in cropping systems: a system approach illustrated on chickpea in India and Europe. In: Karkwal, M.C. (Ed.), *Food Legumes for Nutritional Security and Sustainable Agriculture*. 1, 820-836. Indian Society of Genetics and Plant Breeding (India).
- Wery, J., Adam, M., Deheuvels, O., Delmotte, S., Lamanda, N., Metay, A., Merot, A., Ngo Bieng, M.A., Nicolas, D., Rapidel, B., 2009. Conceptual modelling of a cultivated field: a generic methodology and its use in agronomy. *Farming System Design Symposium*, Monterey (USA) 23-26/08/09.
- Williams J.R., Jones C.A., Kiniry J.R., Spanel D.A., 1989. The EPIC Crop Growth Model. *Transactions of the Asae* 32, 497-511.
- Yin, X., Struik, P.C., Kropff, M.J., 2004. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science* 9, 426-432.
- Yin, X., Van Laar, H.H., 2005. *Crop Systems Dynamics: An ecophysiological model for genotype-by-environment interactions*. Wageningen Academic Publishers, Wageningen, The Netherlands. 168 pp.
- Zahran, H.H., 1999. Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiology and Molecular Biology Reviews* 63, 968.
- Zander, P., Borkowski, N., Hecker, J.M., Uthes, S., Stokstad, G., Rørstad, P.K., Bellochi, G., 2010. Procedure to Identify and Assess Current Activities, SEAMLESS Report no. 50, SEAMLESS integrated project, EU 6th Framework Programme, contract no. 010036-2, www.SEAMLESS-IP.org, 124 pp., ISBN no. 978-90-8585-593-4.
- Zeigler, B., Praehofer, H., Kim, T.G., 2000. *Theory of modeling and simulation: integrating discrete event and continuous complex dynamic systems*. Academic Press, San Diego CA, USA.
- Zeigler, B.P., Marshall, C.Y., 1991. *Object-Oriented Modeling and Discrete-Event Simulation*. *Advances in Computers*. Elsevier, pp.67-114.

APPENDICES

APPENDIX A

Description of the crop component of APES (CEREALS)

A short description of other components is included when relevant for crop growth and development

APPENDIX B

Description of the crop component of APES (LEGUMES)

Description of the new modules and changes in existing ones for MS_{pea}

APPENDIX A

Description of the crop component of APES (CEREALS)

Phenology

Crop development, i.e. the order and rate of appearance of vegetative and reproductive organs, is defined in terms of phenological developmental stage (Eq. 1, DVS) as a function of temperature sum, i.e. cumulative daily effective temperature (T_e , Eq. 2). DVS from 0 to 1 corresponds to the vegetative phase (from emergence to anthesis) and from 1 to 2 to the reproductive phase (from anthesis to physiological maturity).

We added to the original LINTUL, a photoperiodic response (f_p as in LINTUL3) and a vernalization delay (f_v , Eq. 2). Vernalization represents the need for a cold period for flower initiation of the crop. It reduces temperature sum accumulation (T_{sum}) between a low threshold temperature for optimal vernalization and a high threshold temperature for optimal vernalization.

$$(DVS)_0^1 = \frac{(T_{sum})_0^1}{T_{sum_{anthesis}}} \text{ and } (DVS)_1^2 = \frac{(T_{sum})_1^2}{T_{sum_{maturity}}} \quad (1)$$

$$\frac{dT_{sum}}{dt} = T_e \times f_p \times f_v \quad (2)$$

$$T_e = \max(0, T_{average} - T_{base})$$

Leaf area development

The LAI dynamics approach is based on temperature and leaf dry matter production driven by the development stage of the crop (i.e. phenology). During the juvenile phase, LAI expansion is governed by temperature, through its effect on cell division and extension, described via temperature-dependent relative growth rate; this phase thus follows an exponential growth pattern (Eq. 3).

$$\frac{dLAI_{growth}}{dt} = LAI \times R_g \times T_e \quad (3)$$

Following this phase, leaf area expansion is restricted by the supply of dry matter and is calculated by multiplying the simulated rate of increase in leaf dry weight, which is based on the total amount of biomass produced multiplied by a leaf biomass allocation factor, with the specific leaf area (S_{LA} , Eq. 4).

$$\frac{dLAI_{growth}}{dt} = \frac{dW_{leaf}}{dt} \times S_{LA} \quad (4)$$

Finally, leaves senesce either due to (i) self-shading (r_{d-sh}) when LAI reaches a critical value ($LAI_{critical}$) and/or (ii) ageing after anthesis (r_{d-ag}).

$$\frac{dLAI_{sen}}{dt} = -r_d \times LAI \quad (5)$$

with: $r_d = \max(r_{d-ag}, r_{d-sh})$

The r_{d-ag} is a function of the temperature and is specified by the following (x,y) pairs: (-10 , 0.03) , (10 , 0.03) , (15 , 0.04) , (30 , 0.09) , (50 , 0.09), where -10, 10, 15, 30 and 50 are temperatures and the values 0.03, 0.04 and 0.09 are the corresponding relative death rates due to ageing. Death of leaves due to ageing only occurs after anthesis, as indicated by $T_{sum} \geq T_{sum-anthesis}$.

Root development (link with the root distribution component)

Root development comprises two components: root dry matter growth and root distribution. Root dry matter growth is derived from total biomass production multiplied by a biomass allocation factor. The root distribution component estimates the partitioning of fine roots between layers in the soil profile. The number of layers is fixed at the start of the simulation (different from the soil horizons, to which soil characteristics are attributed). Root growth rate is affected according to the current state of soil water and nutrients in a layer. Parameters can be fitted to adjust the water and distance to collar sensitivity so that root profiles of most species can be simulated. It is assumed that the horizontal distribution of crop roots is homogeneous, given that the root distribution is a one dimensional simplification of the system. The model is derived from the Hi-sAFe 3D model (Root Voxel Automaton, Mulia, 2005; Mulia and Dupraz, 2006).

Dry matter production

The approach currently implemented for dry matter production is based on a linear relationship between accumulated intercepted photosynthetically active and accumulated biomass. The slope of this linear relation is called radiation use efficiency (RUE , Monteith, 1977) and summarizes the combined effect of photosynthesis and respiration. The daily intercepted photosynthetically active radiation (PAR_{int} , Eq. 6) is assumed to increase with leaf area index (LAI) following a negative exponential function of LAI that is characterized by a crop specific radiation extinction coefficient (calculated within the light interception component). The photosynthetically active radiation is about half of the daily incoming radiation (R_{dr} , $MJ\ m^{-2}\ d^{-1}$). Equation 7 described the daily rate of dry matter accumulation (NPP , $g\ m^{-2}\ d^{-1}$).

$$PAR_{int} = R_{dr} \times 0.5 \times (1 - e^{-k \times LAI}) \quad (6)$$

$$NPP = RUE \times PAR_{int} \quad (7)$$

Dry matter allocation

The partitioning of dry matter among the different organs is described via a set of distribution factors (default allocation tables, Boons-Prins et al., 1993) that are defined as a function of development stage (Eq. 8). Before anthesis, most dry matter is allocated to the roots, leaves and stems. After anthesis, dry matter accumulates mostly in the storage organs. It is assumed that severe water stress will lead to increased allocation of dry matter to the roots, at the expense of allocation to the shoots.

$$\frac{dW_i}{dt} = P_{c_i} NPP \quad (8)$$

where $\frac{dW_i}{dt}$ and Pc_i are the rate of growth ($\text{g m}^{-2} \text{d}^{-1}$) of organ i and the biomass partitioning factor to organ i ($\text{g organ-}i \text{ g}^{-1} \text{ biomass}$), respectively.

Further we added to LINTUL3, reallocation of dry matter from leaves and stems to storage organs that starts after anthesis for cereals and could be of importance for other crops as well (Eq. 9).

$$\frac{dW_{i\text{realloc}}}{dt} = \frac{dDVS}{dt} \times (W_i - ((1 - \text{frac}W_{i\text{realloc}}) \times W_{i\text{MAX}})) \quad (9)$$

with i = leaves or stems, and

$$\begin{aligned} \text{frac}W_{\text{leafrealloc}} &= \frac{W_{\text{leafmax}} - W_{\text{dead Leaf_endcycle}}}{W_{\text{leafmax}}} \\ \text{frac}W_{\text{stemsrealloc}} &= \frac{W_{\text{stemsmax}} - W_{\text{stems_endcycle}}}{W_{\text{stemsmax}}} \end{aligned}$$

Water dynamics

Potential plant transpiration

The Penman equation (Eq. 10, Penman, 1956) is used to calculate potential evapotranspiration (from the CLIMA component: weather component included in APES) and the leaf area index is used to calculate partitioning between potential evaporation and potential transpiration.

$$T_{\text{pot}} = ET_{\text{ref}} \times (1 - e^{(-0.5 \times LAI)}) \quad (10)$$

with T_{pot} potential transpiration in mm d^{-1} , ET_{ref} the reference evapotranspiration in mm d^{-1} , taking into account the soil and crop albedo, and LAI the leaf area index in $\text{m}^2 \text{ leaf m}^{-2}$ surface area. The 0.5 corresponds to the average extinction coefficient for visible and near infrared radiation, as total radiation (rather than PAR) contributes to evapotranspiration.

Water capture (water uptake component)

From crop water demand (i.e. potential transpiration) and root length density (from the root distribution component), the model computes a local water demand in each soil layer, proportional to the fraction of roots present in the layer. Available water for the crop is the difference between current soil water content and soil water content at wilting point.

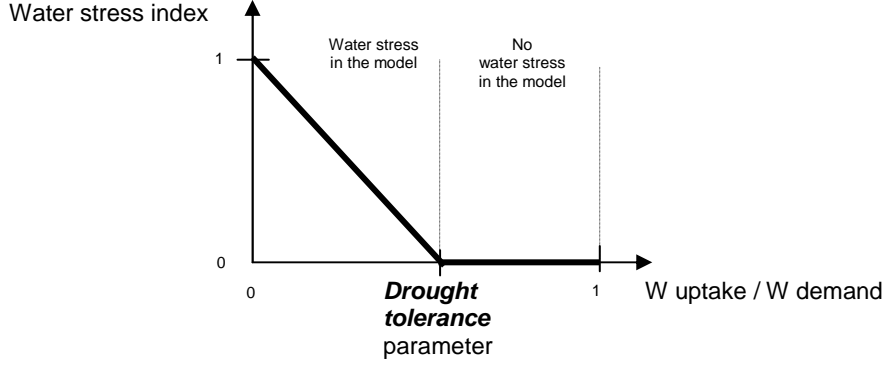
If in a particular layer available water exceeds local water demand, water uptake from that layer equals water demand. If available water is lower than the local water demand, the crop takes up all the available water from this layer and the unsatisfied demand is distributed among the other soil layers in a second loop. If no more water is available in any rooted layer, water demand (i.e. potential transpiration) is not met, and the crop experiences water stress.

Water stress

The water stress index ($water_{\text{stress}}$) is calculated as a function of actual transpiration (T_a) and potential crop transpiration (T_{pot}). It varies from 0 (no water stress) to 1 (maximum water stress). Water stress affects mainly daily growth via an adjustment of the light use efficiency, but only above a given threshold level. This adjustment is operationalized through a

genotypic parameter (the drought tolerance factor, D_t) whose value represents the ability of the species to resist drought. It is defined as the specific value of the ratio of water uptake to water demand, below which the plant starts to experience water stress.

$$water_{stress} = 1 - \min(1, \frac{T_a}{T_{pot}} \times \frac{1}{D_t}) \quad (11)$$



Water stress affects three processes:

1. leaf development during leaf area expansion in the juvenile stage

$$\frac{dLAI_{growth}}{dt} = LAI \times R_g \times T_{eff} \times (1 - water_{stress}) \quad (12)$$

2. total crop growth:

$$\frac{dW}{dt} = LUE \times PAR_{int} \times (1 - water_{stress}) \quad (13)$$

3. allocation to leaves: Dry matter partitioning changes in favour of root growth during the vegetative phase (Munns and Pearson, 1974) when the ratio of actual to potential transpiration falls below 0.5 (Van Keulen et al., 1987).

$$(P_{root})_{ws} = P_{root} \times \max(1, \frac{1}{(1 - water_{stress}) + 0.5}) \quad (14)$$

Crop nitrogen dynamics (Shibu et al., 2010)

Crop nitrogen capture

Based on the crop nitrogen demand and the root length density (from the root distribution component), the model computes a nitrogen demand in each soil layer proportional to the fraction of roots present in the layer. Nitrogen available for the crop is the difference between total soil nitrogen and microbial nitrogen demand (derived from Corbeels et al., 2005).

Crop nitrogen demand

The rate of change in nitrogen ($\frac{dN_{content}}{dt}$, $g\ m^{-2}\ d^{-1}$) in each organ (with i referring to leaves, stems and roots) is calculated as (Eq. 15):

$$(\frac{dN_{content}}{dt})_i = (\frac{dN_{uptake}}{dt})_i - (\frac{dN_{translocated}}{dt})_i - (\frac{dN_{loss}}{dt})_i \quad (15)$$

The rate of N uptake ($\frac{dN_{\text{uptake}}}{dt}$, g m⁻² d⁻¹) for each organ is the result of total nitrogen taken up by the crop, partitioned in proportion to the demand of the organ (Eq. 16). Nitrogen uptake ceases at anthesis, because nitrogen content in the vegetative parts hardly increases after anthesis (Sinclair and Amir, 1992).

$$\left(\frac{dN_{\text{uptake}}}{dt}\right)_i = \frac{(N_{\text{demand}})_i}{(N_{\text{demand}})_{\text{total}}} \times \frac{dN_{\text{uptake}}}{dt} \quad (16)$$

N demand of individual organs, $(N_{\text{demand}})_i$ (g m⁻² d⁻¹) with i for leaves, stems, roots and storage organs, is calculated as the difference between potential and actual organ nitrogen contents. Potential nitrogen content is derived from the maximum nitrogen concentration of an organ ($N_{\text{max},i}$, g (gDM)⁻¹) defined as a function of crop development stage (Eq. 17). The parameter $\Delta_{\text{UptakeMassflow}}$ defines the number of days needed for N uptake from the soil to satisfy the demand of each organ. We assume that 9 days ($\tau = 3$) are necessary to approximately complete this process.

Total crop nitrogen demand equals the sum of the nitrogen demands of the individual organs (excluding storage organs, for which nitrogen demand is met by translocation after anthesis from the other organs, i.e. roots, stems and leaves).

$$(N_{\text{demand}})_i = \frac{(N_{\text{max},i} \times W_i - AN_i)}{\Delta_{\text{UptakeMassflow}}} \quad (17)$$

Nitrogen demand of the grains (storage organs) is met exclusively by translocation from leaves, stems, and roots, as soon as grain growth starts.

Total nitrogen available for translocation in the crop equals total nitrogen content of the organs (N_{content} , g m⁻²) minus their residual non-transferable nitrogen contents, i.e. the nitrogen incorporated in structural crop components. The rate of translocation of N from each organ i to the grain ($\frac{dN_{\text{translocated}}}{dt}$, g m⁻² d⁻¹) corresponds to N demand of the grain ($\frac{dN_{\text{demand}}}{dt}$)_{grain}, g m⁻² d⁻¹) multiplied by the translocated nitrogen of each organ (Eq. 18).

$$\left(\frac{dN_{\text{translocated}}}{dt}\right)_i = \left(\frac{dN_{\text{demand}}}{dt}\right)_{\text{grain}} \times \frac{(N_{\text{content}})_i - [W_i \times (N_{\text{conc not translocable}})_i]}{(N_{\text{content}})_{\text{total}}} \quad (18)$$

N loss corresponds to the nitrogen lost due to death of the organ (Eq. 19, here, only leaves and roots)

$$\left(\frac{dN_{\text{loss}}}{dt}\right)_i = (N_{\text{conc not translocable}})_i \times \left(\frac{dW_{\text{dead}}}{dt}\right)_i \quad (19)$$

Nitrogen stress

To simulate nitrogen-limited crop growth, the rate of dry matter accumulation is reduced when crop nitrogen concentration falls below a critical value. The nitrogen status of the crop is assessed by the nitrogen nutrition index (NNI , Eq. 20), defined as actual N concentration above the residual divided by a critical N concentration above the residual. To calculate this index for the crop as a whole, individual plant organs are considered.

$$NNI = \frac{N_{\text{actual total}} - N_{\text{residual}}}{N_{\text{critical}} - N_{\text{residual}}} \quad (20)$$

Critical crop nitrogen concentration (Eq. 21), the lower limit of canopy nitrogen concentration in leaves and stems required for unrestricted growth, has been set to half the maximum nitrogen concentration (Jamieson et al., 1998).

$$N_{\text{critical}} = \sum_{i=1}^n N_{\text{max},i} \times (N_{\text{optimal}})_{\text{fraction}} \quad (21)$$

$$N_{\text{stress}} = 1 - NNI \quad (N_{\text{stress}} = 1 \text{ being maximum stress})$$

Nitrogen stress affects three processes*:

1. leaf development

a. leaf area expansion in the juvenile stage:

$$\frac{dLAI_{\text{growth}}}{dt} = LAI \times R_g \times T_{\text{eff}} \times NNI \times (1 - water_{\text{stress}}) \quad (22)$$

b. Leaves senescence : If leaf nitrogen concentration per unit leaf area decreases, as nitrogen is translocated from the oldest leaves at the bottom of the canopy to newly formed leaves or to grains, leaves die:

$$\left(\frac{dW}{dt}\right)_{\text{loss}_{\text{ns}}} = W_{\text{lv}} \times R_{(\text{death})_{\text{ns}}} \times N_{\text{stress}} \quad (23)$$

$$\left(\frac{dLAI}{dt}\right)_{\text{loss}_{\text{ns}}} = LAI \times R_{(\text{death})_{\text{ns}}} \times N_{\text{stress}} \quad (24)$$

2. total crop growth:

$$\frac{dW}{dt} = RUE \times PAR_{\text{int}} \times NNI \times (1 - water_{\text{stress}}) \quad (25)$$

3. allocation to leaves: It is assumed that severe nitrogen stress will lead to decreased allocation of dry matter to leaves

$$(P_{\text{leaves}})_{\text{ns}} = P_{\text{leaves}} \times e^{-N_{\text{stress}}} \quad (26)$$

*In this particular modeling solution, we assumed that the nitrogen and water stresses can be multiplied rather than considering the most limiting factor of the two as the main limiting factor (as in Shibu et al., 2010). To our knowledge, no complete agreement on how to model the interaction of the two stresses is reached. Therefore, we are aware that such assumption could be discussed and a new modeling solution could be created to test alternative hypotheses.

APPENDIX B

Description of the crop component of APES: NEW MODULES FOR LEGUMES

Phenology

As for cereals, phenological development of pea is mainly driven by temperature, but co-determined by water stress (Ney et al., 1994). The indeterminate behavior of the pea is affected by water stress: it was assumed that the length of the flowering period is reduced under water stress (threshold value: ws_{par}), considering the beginning of flowering as $T_{sum_{beginningflower}}$ and the end of flowering as $T_{sum_{anthesisws}}$.

$$T_{sum} = \int T_e dt \quad (27)$$

$$(DVS)_0^1 = \frac{T_{sum}}{T_{sum_{anthesisws}}} \text{ and } (DVS)_1^2 = \frac{T_{sum}}{T_{sum_{maturity}}} \quad (28)$$

$$\begin{aligned} T_{sum_{anthesisws}} &= T_{sum_{anthesis}} \times f_{ws} \\ \text{if}(ws \geq ws_{par}) f_{ws} &= \frac{T_{sum_{anthesis}} - T_{sum_{beginningflower}}}{T_{sum_{anthesis}}} \\ \text{else } f_{ws} &= 1 \end{aligned} \quad (29)$$

Nitrogen capture

We assumed that pea has a lower potential of nitrogen uptake from the soil than wheat which may be related to a lower fine root density in the surface soil layers (Gregory, 1998) and lower activity of nitrate transporters (Wery, 1996). Therefore, we adjusted the nitrogen demand of the crop ($N_{demand_{plant}}$ for each organ i referring to leaves, stems and roots to define total nitrogen demand from the soil, by a reduction coefficient k . We assumed an increase in k under water stress, reflecting the higher nitrogen uptake from the soil by the legumes under dry conditions (Mahieu et al., 2009).

$$(N_{demand})_i = \frac{(N_{max,i} \times W_i - AN_i)}{\Delta U_{ptakeMassflow}} \quad (30)$$

$$N_{demand_{plant}} = \sum_{i=1}^n N_{demand_i} \quad (31)$$

$$N_{demand_{soil}} = k \times N_{demand_{plant}} \quad (32)$$

Daily nitrogen fixation (N_{fixed}) is defined as the difference between daily crop nitrogen demand and daily nitrogen uptake from the soil. We assumed that N_2 fixation can fully meet the nitrogen requirements of the legumes if nitrogen uptake from the soil is insufficient.

$$\text{if}(0.3 \leq DVS \leq 1.5)$$

$$N_{fixed} = N_{demand_{plant}} - N_{demand_{soil}} \quad (33)$$



SUMMARY /SAMENVATTING/ RÉSUMÉ

GLOSSARY

ACKNOWLEDGEMENTS

CV and LIST OF PUBLICATIONS

PE&RC EDUCATION STATEMENT

FUNDING

Crop growth models are, increasingly, applied for modelling crop growth and development under a wide range of climatic and soil conditions, under various management types (e.g. intensive vs. extensive agriculture) and at various scales (from field to globe). In response to these diverse uses of models, new software engineering techniques have recently been adopted in the crop modelling field to construct crop modelling frameworks that create flexibility and modularity in crop growth modelling, thus offering simulation options for different processes relevant to agricultural production systems. Such frameworks facilitate adaptation of the model structure to the objective of the simulation, to the production situation (different environments and management practices), to the spatial scale and to data availability. The study presented in this thesis gives insights in the ways to deal with the flexibility of these frameworks. It shows that one of the main issues for handling flexibility in crop modelling is to conciliate crop physiology and agronomy with software engineering. Combining these different disciplines involves: (i) the gathering of different modules describing crop growth and development processes, to constitute a library of alternative modules, and (ii) the development of a set of approaches to facilitate the assembly of the appropriate modules for a specific objective, leading to the construction of a library of assemblies of modules (i.e. crop growth models). The main outcome of this thesis is a framework providing a set of approaches to build models in a guided way (i.e. to select the appropriate modules for a specific model structure aimed at realizing specific objectives).

The first requirement in building the framework is to create the libraries. Construction of these libraries requires the adoption of new software engineering techniques. Although comparable approaches have been described in the literature, the novelty of the approach described in *Chapter 2* resides in the explicit description of the relation between the software design that provides modularity to the framework and the definition of criteria used in selecting the modules to assemble. This relation is achieved through the use of two main design patterns (i.e. general reusable solutions to a commonly occurring problem in software design). The strategy design pattern enables the construction of a library of modules, while the abstract factory design pattern enables the creation of a library of crop growth simulators (various assemblies of the modules), each with a specific model structure (i.e. a modelling solution, MS). These model structures are defined on the basis of explicit criteria. Application of these design patterns has been operationalized in the decision-making software CROSPAL (CROp Simulator: Picking and Assembling Libraries).

The use of CROSPAL is illustrated for the modelling of crop phenology (*Chapter 2*). It exemplifies the collaboration between a crop expert (i.e. a potential user of CROSPAL) wanting to establish the right timing for a pesticide application on wheat and a

software engineer. This collaboration resulted in the inclusion of an alternative module for leaf appearance rate modelling within CROSPAL, thus expanding the library of modules. It resulted in the construction of a new crop growth simulator (MS) that allows simulation of the number of leaves - essential information for the expert to determine when to apply the pesticide. I concluded that the software design presented in *Chapter 2* helps to deal with the necessary flexibility in crop modelling, but that further development and evaluation of this design is required to be able to extend its range of application to more processes currently addressed in crop modelling (e.g. resource competition, nitrogen stress effects on crop growth).

To define the modelling solutions (i.e. the library of crop simulators), there is a need to explicitly formulate the validity domains and the underlying hypotheses of the different modelling approaches of essential crop growth and development processes (i.e. the library of modules) integrated in the framework. *Chapter 3* addresses these issues and presents a systematic methodology to consistently describe the assumptions underlying each module. This methodology is elaborated in a protocol for a systematic selection and evaluation of a crop growth model structure (i.e. the modelling solution, MS). The protocol comprises three main steps: (i) MS selection, (ii) MS calibration, and (iii) MS evaluation and improvement. In particular, steps i and iii of the protocol rely on a thorough analysis of the different modelling approaches (modules), resulting in the selection of specific modules (step i), and on the procedure to assemble the modules to create a new modelling solution adapted to the simulation objective (i.e. step iii: MS improvement).

By applying the proposed protocol for the simulation of crop growth in response to nitrogen management (*Chapter 3*), we identified the importance of the documentation of (i) the different modules included in the framework and (ii) the modelling decision-making process to clarify the sources of uncertainties, associated with the different modules. For instance, in our case study, we explicitly formulated the assumptions underlying the module describing the nitrogen mineralisation process, and particularly we emphasized the importance of the contribution of soil microbial activity to the total nitrogen uptake from the soil under study (a soil with a high organic matter content). We also demonstrated that the use of this protocol (i) promotes the inclusion of different modelling approaches in one tool (instead of having a plethora of individual crop growth models) and (ii) facilitates reuse of (parts of) a model. Indeed in *Chapter 3*, we could easily reuse an existing soil nitrogen module that included an explicit description of soil microbial activity, thus creating a new MS more appropriate for our simulation objective.

By collecting different modelling approaches (i.e. modules) for crop growth processes, we logically found various levels of detail incorporated in crop growth models, reflecting the models' objectives. Good modelling practice dictates to keep models as simple as possible, but enough detail should be incorporated to describe the major processes that determine the behaviour of the system to be modelled. To investigate the effect of the level of detail incorporated in process-based crop growth models, we used the principles of our framework to simulate potential yields under a wide range of climatic conditions (*Chapter 4*). We focused our analysis on the key crop growth processes of leaf area expansion and biomass production under different regimes of temperature and radiation. Our analysis showed that by using a constant radiation use efficiency (RUE) value under a wide range of climatic conditions, the description of the process of biomass production may be over-simplified, as the effects of high temperatures and high radiation intensities on the value of RUE are thus neglected. The effect of temperature and radiation should be considered not only on a daily basis (by using the detailed photosynthesis approach according to Farquhar), but also on a seasonal basis, by describing the value of RUE as a function of temperature and radiation (when using a less detailed description of the process of the biomass production). We derived a linear relation between RUE calculated over the growing season on the one hand, and the average seasonal temperature and seasonal accumulated radiation on the other hand. This relation can be used for further development of simple crop growth models to be applied at global scale.

Furthermore, the two different approaches for the calculation of light interception studied (both described in terms of leaf area development: in one approach a function of temperature and leaf dry matter supply, itself a function of phenology, and in the other approach a function of phenology only) resulted in significant differences in simulated yield (*Chapter 4*). We also found that the light interception approaches better explained the differences in yield sensitivity to climatic variability than the biomass production approaches. After clarifying the assumptions underlying identification of the parameter that represents the onset of leaf senescence in both light interception approaches, we concluded that a better understanding of the whole leaf senescence process is still needed. Misrepresentation of the leaf area index dynamics leads to significant over- or under-estimation of crop yields. These conclusions stress that for applications of crop growth models under a wide range of climatic conditions, particular attention should be given to the choice of the light interception description.

The selection of the crop growth and development modules is also related to the crop type considered (e.g. cereals vs. legumes; annuals vs. perennials; temperate vs.

tropical). Modification of an existing modelling solution (i.e. a crop model) through addition, removal and/or modification of modules can be facilitated by the use of a crop modelling framework. In *Chapter 5*, we explained a procedure to develop a new crop model (i.e. re-assemble) from an existing one, using an existing crop modelling framework (APES, Agricultural Production and Externalities Simulator) and crop physiological knowledge. The success of the changes in the modelling solution (i.e. model structure) for a new crop relies on a stepwise approach, supporting the exchange of concepts among the three disciplines involved in the development of crop models (i.e. agronomy, mathematics and software engineering).

We illustrated the approach with a case study in which a wheat crop growth model was transformed into a pea crop growth model (*Chapter 5*). The changes can be introduced at two levels. First, the modules can be modified, resulting in new modules to include in the library of modules. These modifications comprise: (i) changes in the values of parameters, and/or (ii) changes in equations. For instance, in our case study, we assumed the pea crop to have a lower potential for nitrogen uptake than wheat. Thus, we modified the crop nitrogen demand module, initially developed for wheat, by including a new equation to limit crop nitrogen uptake from the soil. Second, the overall modelling solution can be changed. In our case study, we created two new modules that were added to the initial modelling solution, namely a phenology module for indeterminate crops and a nitrogen fixation module. We also removed the nitrogen stress module from the initial modelling solution. We concluded that the approach to re-assemble modules into a new crop growth model by using a crop modelling framework leads to: (i) integration of knowledge from different disciplines around a modelling objective, (ii) combination of new (expert) knowledge with existing models without ‘re-inventing the wheel’, and (iii) efficient communication with the user of the tool. We demonstrated that through the use of the crop modelling framework and of conceptual modelling (using visual tools), we could easily transform one modelling solution into another, building on the combined expertise of the different disciplines.

Overall, this thesis shows the main benefits, but also the limitations, of the development and use of crop modelling frameworks. I have illustrated that to include modularity and flexibility into crop modelling, we need to: (i) define the system in terms of the basic crop growth and development processes, (ii) provide different modelling approaches (modules) for each process, (iii) explicitly formulate expert knowledge to ensure consistency in the selection of modelling approaches for a specific application, and (iv) use design patterns (i.e. strategy and abstract factory

design patterns). Further, I discussed that a crop modelling framework can be used for: (i) synthesis and integration of results of fundamental research, (ii) strategic planning to evaluate policies, and/or (iii) assessment of management practices. Therefore, there is a need to clearly identify the objectives of all the different partners involved, before embarking on the development of the new tool. In addition, as communication between software engineers and users of the framework becomes more and more difficult because of specialisation, I stressed that the crop modeller should act as an intermediary between these two groups. Finally, I argued that the use of a crop modelling framework can contribute to further development of crop modelling (*Chapter 6*). Not only can it support the proper reuse of models for integrated studies, but it can also promote better understanding of crop physiology because it makes testing of alternative hypotheses easier. Combination of software engineering techniques and principles of systems analysis supports selection of crop models that match the simulation objectives of the users. I also emphasized the need for better documentation of modules and of the process of decision-making in modelling. These conclusions reinforce the continuing importance of the principles of systems analysis for the field of crop modelling, while making use of the latest advances in software engineering techniques: “crop modelling: from conceptual modelling to software engineering, and back”.

Gewasgroeimodellen worden meer en meer toegepast voor het modelleren van gewasgroei en -ontwikkeling onder verschillende klimaat- en bodemomstandigheden, bij verschillende beheerstypen (bv. intensieve vs. extensieve landbouw) en op verschillende schaalniveaus (van een perceel tot de gehele aarde). Als antwoord op deze diversiteit in het gebruik van modellen is men bij het modelleren van gewassen de laatste tijd nieuwe programmatuurkundige technieken (software engineering techniques) gaan toepassen om raamwerken voor gewasmodellering te construeren die flexibiliteit en modulariteit aan het modelleren van gewasgroei geven. Daarmee wordt het mogelijk verschillende processen die van belang zijn in landbouwkundige productiesystemen op verschillende manieren te simuleren. Dit soort raamwerken vergemakkelijkt het aanpassen van de modelstructuur aan het doel van de simulatie, aan de productieomstandigheden (verschillende milieus en beheersmaatregelen), aan de ruimtelijke schaal en aan de beschikbaarheid van gegevens. De studie die in dit proefschrift wordt gepresenteerd verschaft inzichten in de manier waarop de modelleerflexibiliteit die deze raamwerken bieden kan worden gebruikt. Hierbij blijkt het op één lijn krijgen van gewasfysiologie en agronomie enerzijds, en software engineering anderzijds één van de voornaamste opgaven te zijn. Het combineren van deze verschillende disciplines brengt met zich mee (i) het verzamelen van verschillende modules die gewasgroei en -ontwikkelingsprocessen beschrijven in een bibliotheek van alternatieve modules, en (ii) het ontwikkelen van werkwijzen om de juiste modules bijeen te brengen voor een specifieke doelstelling, om zo te komen tot de opbouw van een bibliotheek van samenstellen van modules (i.e. gewasgroeimodellen). Het voornaamste resultaat van dit proefschrift is een raamwerk dat aanwijzingen verschaft voor het bouwen van modellen, d.w.z. voor het kiezen van de juiste modules voor een specifieke modelstructuur die gericht is op het bereiken van specifieke doelen.

Het eerste dat nodig is voor de bouw van het raamwerk is het opzetten van de bibliotheken, en bij het bouwen van deze bibliotheken dienen nieuwe software engineering technieken te worden toegepast. Hoewel vergelijkbare benaderingen in de literatuur zijn beschreven, is de benadering beschreven in Hoofdstuk 2 vernieuwend vanwege de expliciete beschrijving van de relatie tussen het software-ontwerp dat modulariteit aan het raamwerk geeft en de definitie van de criteria die gebruikt worden bij het selecteren van de te combineren modules. Deze relatie wordt bereikt door het gebruik van twee belangrijke ontwerppatronen (d.w.z., algemene, herbruikbare oplossingen voor een vaak optredend probleem in het ontwerpen van software). Het strategy-ontwerppatroon maakt de opbouw van een bibliotheek van modules mogelijk, terwijl het abstract factory-ontwerppatroon de opbouw mogelijk

maakt van een bibliotheek van gewasgroeisimulators (verschillende samenstellen van de modules), elk met een specifieke modelstructuur (d.w.z., een modelleeroplossing (modelling solution, MS)). Deze modelstructuren zijn gedefinieerd op basis van expliciete criteria. Toepassing van deze ontwerp patronen is geoperationaliseerd in het beslissingsprogramma CROSPAL (CROp Simulator: Picking and Assembling Libraries).

Het gebruik van CROSPAL wordt geïllustreerd aan de hand van het modelleren van gewasfenologie (Hoofdstuk 2). De samenwerking tussen een gewasdeskundige (i.e. een potentiële gebruiker van CROSPAL) die de juiste timing van een pesticidebespuiting van tarwe wil vaststellen, en een software engineer wordt daarbij als voorbeeld gebruikt. Deze samenwerking resulteerde in het opnemen van een alternatieve module voor het modelleren van de bladverschijningssnelheid binnen CROSPAL, waardoor de bibliotheek van modules werd uitgebreid. Het resulteerde in de bouw van een nieuwe gewasgroeisimulator (MS) waarmee het aantal bladeren – essentiële informatie voor de deskundige om het bespuitingstijdstip vast te stellen – kan worden gesimuleerd. Ik concludeerde dat het software-ontwerp gepresenteerd in Hoofdstuk 2 helpt bij het oplossen van het probleem van de noodzakelijke flexibiliteit in gewasmodellering, maar dat dit ontwerp verder ontwikkeld en geëvalueerd moet worden om het toepassingsbereik tot meer processen waarmee de gewasmodellering zich tegenwoordig bezig houdt (bv. concurrentie om hulpbronnen, effecten van stikstofstress op gewasgroei) te kunnen uitbreiden.

Om de modelleeroplossingen (i.e. de bibliotheek van gewassimulators) te definiëren, dienen de geldigheidsdomeinen van essentiële processen van gewasgroei en -ontwikkeling (i.e. de bibliotheek van modules) welke in het raamwerk zijn geïntegreerd, en de hypothesen die daaraan ten grondslag liggen, expliciet te worden geformuleerd. In Hoofdstuk 3 wordt hierop ingegaan en wordt een systematische methodologie gepresenteerd om de aannames die aan elke module ten grondslag liggen consistent te beschrijven. Deze methodologie is uitgewerkt tot een protocol voor een systematische selectie en evaluatie van een structuur van een gewasgroeimodel (i.e. de modelleeroplossing (modelling solution (MS))). Het protocol bestaat uit drie hoofdstappen: (i) selectie van de MS, (ii) kalibratie van de MS, en (iii) evaluatie en verbetering van de MS. In het bijzonder stappen (i) en (iii) van het protocol steunen op een grondige analyse van de verschillende procesbeschrijvingen (modules), hetgeen resulteert in de selectie van specifieke modules (stap i), en op de procedure voor het combineren van de modules om te komen tot een nieuwe

modelleeroplossing die toegesneden is op het doel van de simulatie (i.e. stap (iii): verbetering van de MS).

Door het toepassen van het voorgestelde protocol op de simulatie van gewasgroei in response op stikstofbeheer (Hoofdstuk 3) werd het belang duidelijk van het documenteren van: (i) de verschillende modules binnen het raamwerk en (ii) het proces van het nemen van modelleerbeslissingen, om duidelijkheid te verkrijgen over de bronnen van onzekerheden die met de verschillende modules samenhangen. Wij hebben bijvoorbeeld in onze casestudy expliciet de aannames geformuleerd die ten grondslag liggen aan de module die het proces van stikstofmineralisatie beschrijft en in het bijzonder hebben we het belang benadrukt van de bijdrage van de microbiële activiteit in de bodem aan de totale stikstofopname uit die bodem (een bodem met een hoog organisch stofgehalte). We hebben ook aangetoond dat het gebruik van dit protocol: (i) het bijeenbrengen bevordert van verschillende modules in één stuk gereedschap (in plaats van dat we een overvloed aan individuele gewasgroeimodellen krijgen), en (ii) het hergebruik van (delen van) een model gemakkelijker maakt. Wij konden in Hoofdstuk 3 inderdaad een bestaande bodemstikstofmodule die een expliciete beschrijving van microbiële activiteit in de bodem bevatte, gemakkelijk hergebruiken om een nieuwe MS te creëren die beter geschikt was voor ons simulatiedoel.

Door het bijeenbrengen van verschillende procesbeschrijvingen (i.e. modules) van gewasgroeiprocessen werd het logischerwijs duidelijk dat gewasgroeimodellen verschillende niveaus van gedetailleerdheid bevatten, waarbij het niveau een afspiegeling is van het doel van het model. Volgens de principes van goed modelleren (Good Modelling Practices) dient men modellen zo eenvoudig mogelijk te houden, maar ze moeten gedetailleerd genoeg zijn om de voornaamste processen te kunnen beschrijven die het gedrag van het te modelleren systeem bepalen. Om het effect van de mate van detail in de beschrijving van processen in op procesbeschrijvingen gebaseerde gewasgroeimodellen te onderzoeken, hebben wij de principes van ons raamwerk gebruikt om potentiële opbrengsten onder zeer verschillende klimaatomstandigheden te simuleren (Hoofdstuk 4). Wij hebben onze analyse toegespitst op sleutelprocessen bij gewasgroei, namelijk de groei van bladoppervlak en de biomassaproductie onder verschillende temperatuurs- en stralingsomstandigheden. Onze analyse liet zien dat het gebruik van een constante waarde voor de stralingsbenuttingsefficiëntie (radiation use efficiency (RUE)) onder zeer verschillende klimaatomstandigheden voor de beschrijving van het proces van biomassaproductie te eenvoudig kan zijn, omdat dan de effecten van hoge temperatuur en hoge stralingsintensiteit op de waarde van RUE worden verwaarloosd. Het effect van temperatuur en straling dient niet alleen te worden beschouwd op dag-basis (wanneer

de gedetailleerde fotosynthesebenadering volgens Farquhar wordt gebruikt), maar ook op seizoensbasis, door de waarde van RUE te beschrijven als functie van temperatuur en straling (wanneer een minder gedetailleerde beschrijving van het proces van biomassaproductie wordt gebruikt). Wij hebben een lineaire relatie afgeleid tussen de over het groeiseizoen berekende RUE enerzijds, en de gemiddelde seizoenstemperatuur en de over het seizoen geaccumuleerde straling anderzijds. Deze relatie kan worden gebruikt voor het verder ontwikkelen van eenvoudige gewasgroeimodellen voor toepassingen op wereldschaal.

Voorts resulteerden de twee benaderingen voor de berekening van de lichtonderschepping die werden bestudeerd (beide beschreven in termen van bladoppervlakte-ontwikkeling: de ene als een functie van temperatuur en aanbod van drogestof voor het blad (die zelf een functie is van de fenologie), de andere alleen als een functie van de fenologie alleen) in significante verschillen in gesimuleerde opbrengst (Hoofdstuk 4). Ook vonden we dat de benaderingen via de lichtonderschepping de gevoeligheid van opbrengsten voor klimaatvariabiliteit beter verklaarden dan de benaderingen via de biomassaproductie. Nadat we de aannames opgehelderd hadden die ten grondslag liggen aan de definitie van de parameter die het begin van de bladveroudering vertegenwoordigt in beide lichtonderscheppingsbenaderingen, concludeerden we dat een beter begrip van het gehele bladverouderingsproces nodig is. Een onjuiste berekening van de dynamiek van de bladoppervlakte-index (leaf area index (LAI)) leidt tot belangrijke over- of onderschatting van de gewasopbrengsten. Deze conclusies benadrukken dat bij het gebruik van gewasgroeimodellen voor een breed scala van klimaatomstandigheden speciale aandacht dient te worden gegeven aan de keuze van de beschrijving van de lichtonderschepping.

De selectie van modules voor gewasgroei en –ontwikkeling hangt ook samen met het type gewas dat beschouwd wordt (granen vs. peulgewassen; eenjarige vs. overblijvende; gematigde vs. tropische). Veranderingen in een bestaande modelleeroplossing (i.e. gewasmodel) door toevoeging, verwijdering en/of wijziging van modules kunnen worden ondersteund door het gebruik van een gewasmodelleringsraamwerk. In Hoofdstuk 5 hebben we een procedure uitgelegd waarmee een nieuw gewasmodel kan worden ontwikkeld (d.w.z., opnieuw samengesteld) uitgaande van een bestaand model door gebruik te maken van een bestaand gewasmodelleringsraamwerk (viz. APES, Agricultural Production and Externalities Simulator) en van kennis van de gewasfysiologie. Het succes van de veranderingen in de modelleeroplossing (i.e. in de modelstructuur) ten behoeve van

een nieuw gewas steunt op een stapsgewijze benadering die de uitwisseling van concepten tussen de drie vakgebieden die bij de ontwikkeling van gewasmodellen betrokken zijn - landbouwkunde, wiskunde en software engineering - ondersteunt.

We hebben deze benadering geïllustreerd met een casestudy waarin een gewasgroeimodel voor tarwe getransformeerd werd in een model voor erwten (Hoofdstuk 5). De veranderingen kunnen worden aangebracht op twee niveaus. Ten eerste kunnen de modules worden aangepast, wat nieuwe modules oplevert die in de modulenbibliotheek worden opgenomen. De aanpassingen bestaan uit: (i) veranderingen in de waarden van parameters, en/of (ii) veranderingen in vergelijkingen. In onze casestudy hebben we bijvoorbeeld aangenomen dat erwten een lager vermogen tot stikstofopname heeft dan tarwe. Daarom hebben we de module voor stikstofbehoefte, die aanvankelijk voor tarwe ontwikkeld was, aangepast door er een nieuwe vergelijking in op te nemen die de stikstofopname van een gewas uit de bodem beperkt. Ten tweede kan de gehele modelleeroplossing worden veranderd. In onze casestudy hebben we twee nieuwe modules gecreëerd die aan de oorspronkelijk modelleeroplossing werden toegevoegd, namelijk een module voor de fenologie van gewassen met een onbepaalde groeiwijze en een module voor stikstofbinding. Daarnaast hebben we de module voor stikstofstress uit de oorspronkelijke modelleeroplossing verwijderd. We concludeerden dat de benadering van het opnieuw tot één geheel bijeen brengen van bestaande modules tot een nieuw gewasgroeimodel door gebruik te maken van een gewasmodelleringsraamwerk leidt tot: (i) de integratie van kennis uit verschillende vakgebieden rond een modelleringsdoel, (ii) de combinatie van nieuwe kennis (van deskundigen) met bestaande modellen zonder 'het wiel opnieuw uit te vinden', en (iii) een efficiënte communicatie met de gebruiker van dit stuk gereedschap. Wij hebben aangetoond dat door het gebruik van het gewasmodelleringsraamwerk en van conceptuele modellering (gebruik makend van visuele gereedschappen) wij op eenvoudige wijze de ene modelleeroplossing konden transformeren in de andere, bouwend op de gecombineerde expertise uit de verschillende vakgebieden.

Over het geheel genomen toont dit proefschrift de voornaamste voordelen, maar ook de beperkingen, van de ontwikkeling en het gebruik van gewasmodelleringsraamwerken. Ik heb geïllustreerd dat om modulariteit en flexibiliteit in gewasmodellering in te brengen het noodzakelijk is om: (i) het systeem te definiëren in termen van de basisprocessen van gewasgroei en -ontwikkeling, (ii) de beschikking te hebben over verschillende procesbeschrijvingen (modules) voor elk proces, (iii) kennis van deskundigen expliciet te formuleren om consistentie te

waarborgen in de selectie van modules voor een specifieke toepassing, en (iv) ontwerp patronen te gebruiken (i.e. strategy- en abstract factory-ontwerppatronen). Ik heb voorts bediscussieerd dat een gewasmodelleerraamwerk kan worden gebruikt voor: (i) synthese en integratie van resultaten van fundamenteel onderzoek, (ii) strategische planning om beleid te evalueren, en/of (iii) het beoordelen van beheersmaatregelen. Het is daarom noodzakelijk om duidelijk de doelstellingen van alle betrokken partners vast te stellen voordat de ontwikkeling van een nieuw stuk gereedschap ter hand wordt genomen. Tevens heb ik benadrukt dat, vanwege de als gevolg van specialisatie steeds moeilijker wordende communicatie tussen software engineers en gebruikers van het raamwerk, de gewasmodelleur als bemiddelaar tussen deze twee groepen zou moeten optreden. Ten slotte heb ik betoogd dat het gebruik van een gewasmodelleringsraamwerk kan bijdragen aan de verdere ontwikkeling van gewasmodellering (Hoofdstuk 6). Het kan niet alleen het juiste hergebruik van modellen voor geïntegreerde studies ondersteunen, maar het kan ook leiden tot verbeterde inzichten in de gewasfysiologie, omdat het het testen van alternatieve hypothesen vergemakkelijkt. De combinatie van software engineering technieken en principes van systeemanalyse ondersteunt de selectie van gewasmodellen die aan de simulatiedoelen van de gebruikers voldoen. Ik heb ook benadrukt dat de modules en de besluitvormingsprocessen in het modelleren beter gedocumenteerd moeten worden. Deze conclusies ondersteunen het feit dat, terwijl wij gebruik maken van de nieuwste ontwikkelingen op het gebied van software engineering technieken, de principes van de systeemanalyse voor de gewasmodellering voortdurend van belang blijven: “gewasmodellering: van conceptueel modelleren naar software engineering, en terug”.

Les modèles de cultures sont de plus en plus utilisés dans un large éventail de conditions climatiques et pédologiques, pour divers types de systèmes de culture (par exemple, intensif vs. extensif) et à de nombreuses échelles (de la parcelle au globe). En réponse à cette diversité d'utilisation, le domaine de la modélisation des cultures a récemment adopté les pratiques du génie logiciel, créant flexibilité et modularité dans les plateformes de simulation pour offrir des options adaptées à chaque type d'utilisation en terme de processus biophysiques simulés, de variables d'intérêt et de mode de conduite simulés. Ces plateformes facilitent le changement de la structure des modèles en fonction de l'objectif de simulation, des niveaux de productions (différents environnements et pratiques agricoles), de l'échelle et de la disponibilité des données. Le travail présenté dans cette thèse traite de la gestion de la flexibilité de ces plateformes de simulation (*Chapitre 1*). On montre que le principal enjeu pour valoriser cette flexibilité est de réconcilier les connaissances agronomiques avec les pratiques du génie logiciel. La combinaison de ces deux disciplines nécessite (i) de collecter différents modules de croissance et développement de la plante, illustratif de différents processus physiologiques, pour constituer une bibliothèque de modules échangeables, et (ii) d'élaborer un ensemble d'approches qui facilitent l'assemblage des modules appropriés pour un objectif donné de simulation, constituant une bibliothèque d'assemblages de modules. L'aboutissement principal de cette thèse est une plateforme contenant un ensemble d'approches pour construire des modèles de façon guidée (c'est à dire, sélectionner les modules appropriés pour une structure particulière de modèle, correspondant à un objectif de simulation spécifique).

La première condition pour construire la plateforme est de créer les bibliothèques. La construction de ces bibliothèques demande l'adoption des nouvelles techniques de génie logiciel. Bien que des approches comparables aient été décrites dans la littérature, la nouveauté de l'approche décrite dans le *chapitre 2* réside dans la relation entre la conception du logiciel qui fournit modularité à la plateforme de simulation, et la définition de critères qui facilitent l'assemblage des modules. La relation est obtenue par l'utilisation de deux patrons de conception (à savoir des solutions standard pour répondre à des problèmes d'architecture et de conception des logiciels). Le patron de conception *Stratégie* permet de construire une bibliothèque de modules, alors que le patron de conception *Fabrique Abstraite* permet de créer une bibliothèque de simulateurs de cultures (assemblage de modules) ayant une structure de modèle spécifique (aussi appelée "solution de modélisation"), structure déterminée en fonction de critères d'assemblage des modules. Par cette dernière voie de conception de logiciel, nous avons obtenu un logiciel d'aide à la modélisation CROSPAL (CROp Simulator: Picking and Assembling Libraries).

L'utilisation de CROSPAL est illustrée pour la modélisation de la phénologie des cultures (*Chapitre 2*). Avec cet exemple, nous démontrons comment un agronome (utilisateur potentiel de CROSPAL) peut collaborer avec un informaticien, pour définir le moment approprié pour l'application de pesticides sur le blé. Cette collaboration a permis d'inclure dans CROSPAL, un nouveau module représentant une approche de modélisation de la phénologie par le rythme d'apparition des feuilles. Cette inclusion dans la bibliothèque des modules a permis de définir un nouveau modèle de culture simulant le nombre de feuilles, information essentielle pour déterminer quand appliquer le pesticide. L'adoption de cette démarche d'ingénierie logicielle (basée sur des patrons de conception) permet d'analyser la flexibilité de la plateforme de simulation tout en intégrant les connaissances d'experts sur les processus biophysiques simulés. Cependant, nous soulignons aussi qu'il est nécessaire d'utiliser et d'évaluer CROSPAL de manière plus approfondie, notamment sur les questions actuelles de la modélisation des cultures comme, par exemple, la compétition pour les ressources, ou les effets du stress azoté sur la croissance des plantes.

Pour définir correctement les solutions de modélisation (bibliothèque des simulateurs de cultures), il faut formuler explicitement le domaine de validité, et les hypothèses sous-jacentes des différentes approches de modélisations (bibliothèque de modules). Le *chapitre 3* aborde ces questions et présente une méthodologie systématique permettant de définir explicitement les principes représentés dans chaque module. Cette méthodologie est élaborée comme un protocole pour la sélection et l'évaluation systématique de la structure du modèle de culture (solution de modélisation). Le protocole comporte trois étapes principales: (i) la sélection de la solution de modélisation, (ii) la calibration de la solution, et (iii) l'évaluation et l'amélioration de cette solution. Le protocole présenté s'appuie sur (i) une analyse détaillée des différentes approches de modélisation (modules) au cours de l'étape 1, débouchant sur la sélection de modules, et sur (ii) la façon d'assembler les modules pour créer une nouvelle solution de modélisation mieux adaptée à l'objectif de simulation (étape 3: amélioration de la solution).

En utilisant ce protocole pour simuler la réponse de la croissance des cultures à différents modes de gestion de l'azote (*Chapitre 3*), nous avons pu déterminer l'importance de la documentation (i) des différents modules inclus dans la bibliothèque de modules et (ii) du processus de décision de modélisation dans le but de clarifier les sources d'incertitudes liées à la sélection des modules ou aux hypothèses sous-jacentes des modules. Par exemple, dans notre étude de cas, nous avons explicitement formulé les hypothèses du module représentant le processus de

minéralisation de l'azote, et plus particulièrement nous avons identifié l'importance de la contribution de l'activité microbienne du sol dans l'assimilation de l'azote total pour le sol étudié (ayant une haute teneur en matière organique). Nous avons aussi démontré que l'utilisation du protocole, en lien avec une plateforme de simulation telle que celle présentée au chapitre 2, favorise la collecte de différentes approches de modélisation au sein d'un même outil, plutôt que leur dispersion dans une multitude de modèles de cultures. En effet, dans le *chapitre 3*, nous avons pu réutiliser un module existant de l'azote du sol comprenant une description détaillée de l'activité microbienne du sol, pour obtenir une solution de modélisation plus appropriée pour notre objectif de modélisation.

Collectant différentes approches de modélisation (modules) pour les processus de croissance de la culture, nous avons identifié différents niveaux de détail incorporé au sein de ces modèles, reflet de l'objectif initial des créateurs du modèle. Les bonnes pratiques de modélisation imposent de définir un modèle aussi simple que possible, mais de garder assez d'informations sur les processus principaux déterminant le système. Par conséquent, nous avons utilisé les principes présentés précédemment pour étudier l'effet du niveau de détail dans un modèle de cultures développé et testé à l'échelle de la parcelle, pour simuler les rendements à plus grandes échelles spatiales, à savoir l'échelle régionale, nationale voire globale (*Chapitre 4*). Nous avons centré notre analyse sur les processus clés de croissance de la culture, à savoir l'expansion de la surface foliaire et la production de biomasse, sous différents régimes de température et radiation, pour déterminer l'importance relative du niveau de détail pour simuler le rendement potentiel. Nous avons montré que si nous utilisons une valeur constante pour le paramètre d'efficacité d'utilisation de la radiation à grande échelle, nous risquons de trop simplifier le processus de production de biomasse, ignorant les effets de hautes températures et de rayonnements élevés sur ce processus. Non seulement les effets de la température et le rayonnement doivent être considérés sur une base quotidienne (quand on utilise l'approche détaillée de la photosynthèse selon Farquhar), mais aussi sur une base saisonnière lorsque l'on utilise une approche moins détaillée du processus, en ajustant le paramètre RUE pour la température et de rayonnement. Nous avons établi une relation linéaire entre le paramètre RUE calculé sur la saison d'une part, et la température saisonnière moyenne et l'accumulation du rayonnement au cours de la saison, d'autre part. Cette relation peut être utilisée pour la poursuite du développement de modèles de culture simple à l'échelle globale.

De plus, les deux approches étudiées sur l'interception de la lumière (l'une basée sur la température et la production de biomasse de feuilles dépendante de la phénologie, l'autre basée sur la phénologie uniquement) fournissent des différences significatives

en terme de rendements simulés, et permettent de discuter des différences de rendements due à la variabilité climatique (*Chapitre 4*). Après avoir précisé les hypothèses sous-jacentes du paramètre représentant le début de la sénescence des feuilles dans les deux approches simulant la dynamique de l'indice foliaire (c.-à-d. interception de la lumière), nous avons conclu qu'une meilleure compréhension du processus de sénescence des feuilles est encore nécessaire. La mauvaise représentation de la dynamique de l'indice foliaire peut entraîner d'importantes sous ou surestimation des rendements. Ces conclusions soulignent qu'une attention particulière est nécessaire lors de la sélection de l'approche de modélisation définissant l'interception de la lumière pour des applications des modèles de culture à grande échelle, en tenant compte des données disponibles pour le paramétrage et l'évaluation à ces échelles.

La sélection des modules représentant la croissance et le développement des cultures est également liée au type de culture considérée (céréales vs. légumineuses, annuelles vs. pérennes, tempérées vs. tropicales). L'ajout, le retrait ou le changement de modules d'une solution de modélisation existante (modèle de cultures) est facilité par l'utilisation d'une plateforme de simulation. Dans le *chapitre 5*, nous exposons une procédure basée sur une plateforme de simulation et les connaissances d'un expert en écophysiologie des cultures, permettant de développer un nouveau modèle de culture (solution de modélisation) à partir d'une solution déjà existante (ce que nous appelons "ré-assembler" un modèle). Le succès de la modification de la solution de modélisation (structure du modèle) repose sur une approche par étape, permettant d'échanger des notions entre trois disciplines impliquées dans le développement du modèle (agronomie, mathématiques et génie du logiciel). Nous avons illustré cette approche avec une étude de cas construisant un modèle de culture de pois à partir d'un modèle de culture de blé.

Les changements peuvent se produire à deux niveaux. D'abord, les modules peuvent être modifiés, entraînant de nouveaux modules à inclure dans la bibliothèque de modules. Ces modifications comprennent: (i) les changements dans les valeurs des paramètres, et (ii) les changements dans les équations. Par exemple, pour notre étude de cas, nous avons supposé que le pois a un potentiel d'absorption d'azote plus faible que le blé. Ainsi, nous avons modifié le module de la demande en azote des cultures, initialement développé pour le blé en incluant une nouvelle équation pour limiter l'absorption de la plante d'azote du sol. Deuxièmement la solution de modélisation elle-même peut être modifiée. Dans notre étude de cas (*Chapitre 5*) nous avons créé deux nouveaux modules à ajouter à la solution de modélisation initiale, à savoir un module de phénologie de plante indéterminée et un module de fixation de l'azote.

Nous avons également supprimé le module de stress en azote. La structure globale du modèle a été ainsi notablement modifiée par rapport au modèle de culture du blé. Nous avons conclu que l'utilisation de l'approche pour le ré-assemblage d'un nouveau modèle a permis (i) l'intégration des différentes disciplines autour d'un même objectif de modélisation, (ii) une combinaison de nouvelles connaissances avec des modèles existants, sans "réinventer la roue", et (iii) une communication efficace avec l'utilisateur de l'outil aussi bien qu'avec le concepteur du modèle initial et qu'avec l'expert en écophysiologie. Nous avons démontré que l'on pouvait facilement passer d'une solution de modélisation à une autre, grâce à l'utilisation d'une plateforme de simulation et la modélisation conceptuelle (à l'aide d'outils visuels), s'appuyant sur l'expertise de chaque discipline.

Dans l'ensemble, cette thèse a montré les principaux avantages, mais aussi les limites, du développement et de l'utilisation de plateformes de simulation. Il est apparu que pour inclure modularité et flexibilité dans la modélisation des cultures, nous avons besoin (i) de définir le système biophysique en référence aux processus représentatif de la croissance et du développement des cultures, (ii) de fournir différentes approches de modélisation pour chaque processus (modules), (iii) de formuler explicitement les connaissances d'experts pour assurer la cohérence dans la sélection des approches de modélisation pour une application spécifique, et (iv) d'utiliser des patrons de conceptions (Stratégie et Fabrique abstraite) pour fournir la flexibilité à la plateforme de simulation. Nous avons montré qu'une plateforme de simulation de cultures pouvait être utilisée (i) pour la synthèse et l'intégration des connaissances issues de la recherche sur le fonctionnement des cultures, (ii) pour l'évaluation des politiques agricoles, ou (iii) pour l'évaluation de pratiques agricoles. Par conséquent, il est nécessaire de définir clairement l'objectif, entre les différents partenaires concernés, avant de commencer le développement de nouvelle plateforme de simulation ou de la détermination de la solution de modélisation au sein de plateforme existante.

Faisant le constat que les modélisateurs impliqués dans le développement des plateformes et les agronomes impliqués dans l'expérimentation s'éloignent de plus en plus, j'ai souligné l'importance du rôle du modélisateur des cultures en tant que médiateur entre les développeurs et les utilisateurs finaux. Je discute ainsi de la place que peut avoir une plateforme de simulation de cultures dans les développements futurs en modélisation des cultures (*Chapitre 6*). Non seulement ce type de plateforme peut faciliter la réutilisation des modèles existants au plan international, pour des études d'évaluation intégrée des systèmes agricoles, mais elles peuvent favoriser aussi une meilleure compréhension du fonctionnement des cultures en testant diverses hypothèses. La combinaison des techniques de génie logiciel et des principes de

l'analyse des systèmes assiste la sélection de modèles de cultures correspondant aux objectifs de simulation.

Les conclusions de cette thèse confirment le rôle essentiel de l'analyse des systèmes dans le domaine de la modélisation des cultures, en combinaison avec des avancées dans les techniques de génie logiciel: “la modélisation des cultures: de la modélisation conceptuelle des systèmes biophysiques au génie logiciel et réciproquement”.

Whenever the definition was ***appropriate for this thesis***, items in the glossary are defined using/summarizing merriam Webster online (<http://www.merriam-webster.com>) and/or Wikipedia definitions (http://en.wikipedia.org/wiki/Main_Page).

Abstract class: a set of operations which all objects that implement the protocol must support.

Algorithm: a set of coherent equations to solve a specific problem.

APES: Agricultural Production and Externalities Simulator; a crop modelling framework developed within the framework of the SEAMLESS project.

APSIM: Agricultural Production Systems sIMulator; a modular modelling framework that has been developed by APSRU (Agricultural Production Systems Research Unit) in Australia.

Beam: a structural element, i.e. part supporting the framework.

Class: a set of objects having the same behaviour (but typically differing in state), or a template defining such a set.

Component: a piece of software representing plant and/or soil processes that is used to compose a cropping system model (e.g. crop, light interception, water uptake, soil water, or soil C and N components in APES).

Conceptual model: mental model that allows us to understand and simplify the problem

CROSPAL: CROp Simulator: Picking and Assembling Libraries; a decision-making software application for crop modelling.

Design pattern: a general reusable solution to a commonly occurring problem in software design.

Developer: a person responsible for the implementation of the model.

DSSAT: Decision Support System for Agrotechnology Transfer; a collection of independent programs that operate together, crop simulation models are at its centre.

Equation: a mathematical statement.

Factory design pattern: a software design that provides an interface for creating families of related or interdependent objects without specifying the concrete classes.

Flexibility: the quality of being easily adaptable.

Framework: a structure supporting or containing something.

GCROP: a generic crop model template; the APSIM generic crop module designed to implement a process-oriented approach in crop modelling.

Granularity: the extent to which a system is broken down into small parts.

GUI: Graphical User Interface.

Interface: a set of named operations that can be invoked by clients.

Modeller: a person responsible for the actual development of the model.

Modelling solution (MS): a combination of components to construct one effective simulation model.

Modularity: the property of a system to be made up of relatively independent, but interlocking components or parts.

Module: a conceptualization of a specific crop or soil process implemented within a component (e.g. radiation use efficiency or photosynthesis for biomass production).

MS: Modelling Solution; defined within this thesis and in this glossary.

Object: an instance (that is, an actual example) of a class. An object is a self-contained entity that consists of both data and procedures to manipulate the data.

Parameter: a value that determines the characteristics or behaviour of something.

SEAMLESS: System for Environmental and Agricultural Modelling; Linking European Science and Society; European project aiming at the development of a component-based framework for agricultural systems.

Strategy design pattern: a software design that defines a family of algorithms, encapsulates each one, and makes them interchangeable.

User: a person responsible for the use of the model for a specific objective.

Wrapper: a class that serves to mediate access to another.

Synonyms

Basic crop process = module = strategy = modelling approach

Crop model = modelling solution = factory = crop simulator

2006-2010. At first, when I saw this 4-year period on paper, I have to admit that I got a bit scared: being four years in the same place, working on the same thing, how was I going to manage? Well, I quickly felt at home in Wageningen, and these four years went rather fast. This would not have been possible without my roots in northern France (not so far from here) and, most importantly, without the general openness of the Dutch people and the diversity of my supervision team.

Working with this team for four years was a rich experience with an interesting combination of different cultures, personalities and expertise. I have been very fortunate to have had such a supervision team to help me to reach this achievement.

Herman and Jacques, what strikes me the most in both of you is your enthusiasm. There is no doubt how contagious it is and how important it was for my own motivation. I was always “drinking in” your knowledge without any satiety. Herman, your support in my “laaste loodjes” (as you were always saying) was very much appreciated, and clearly helped to get more and more precise in the formulation of my sentences. Jacques, a special thanks for your availability when I was coming to work in Montpellier. It was always a pleasure to come and work with you, and most of all, it helped me to make significant progress, integrating different perspectives.

Peter, I could always pop-up in your office and you will take time to listen to my ups-and-downs. Thanks for tempering some of my moods. Also, I really learnt (at least tried) to pay more attention to details from you. And although I always tend to look at the big picture rather than the particulars, it was always satisfactory to come back to the details and get a hold on the underlying explanations of the phenomena. Thanks for pushing me to do that.

Thanks to Frank who always challenged me and, any time I had a question would dig up some literature that addressed, in one way or another, the problem I was struggling with. Frank, your critical comments on my xth drafts were always exigent, and often pushing me to change direction....but the result was always better science!

Marc Corbeels, this thesis could not be what it is without your scientific contribution and moral support. Besides your always very relevant and challenging comments on the draft of my articles, you would always have a word to cheer me up when it was not easy. Your human skills are as valuable as your scientific ones. Your contribution to this study is, without doubt, precious.

After my supervision team, I want to thank all of my “other” co-authors in articles and conference presentations and “co-workers” of the APES experience for always initiating critical and challenging discussions: Hatem Belhouchette, Eric Casellas, Frits van Evert, Marcello Donatelli, Jacques Eric Bergez, Olivier Thérond, Pablo Tittonnell, Daniel Wallach, Aurélie Perrin, Sebastien Roux, Marco Acutis, Kamal Kansou. Putting together the vision of all of you in one common text was not always easy. BUT, no doubt, without these exchanges of ideas with so many scientists, this work would have been different.

Pytrik, our interactions (especially on the FADN database, and agriculture in Europe) and your tips were always useful for me to progress, and even to decide: “ok, I cannot use these data...well, no problem, I will find something else”. And of course, your positive attitude and always being in for coffees, beers, lunches, sports (great team leader for the We-Day) is undeniably an essential added-value for the social life of the group.

Working within PPS has been a rich experience. Charlotte, Ria, and Gijsbertje, thanks to you, the administrative tasks appeared rather simple. You always came up with ways around and solutions when needed. Thanks for your understanding support during these four years. Ken and Irene, thanks for your lively garden parties. It always felt very “gezellig” to be there. Martin, thanks for arranging a few extra months to my contract, allowing me to get to the defence date with a peaceful mind.

Special thanks to Rik for translating the summary in Dutch (despite my “sloppy English!”); for smoothening my English here and there during these four years; for the cover design; and last but not least for enduring some of my moods (bad and good), even when not always justified.

Finally, I would like to extend my gratitude to all the people that supported me during these four years in one way or another: family, friends and colleagues. Without some of the colleagues in PPS and/or friends met here in Wageningen, in-between-work life would have been less enjoyable. Ilse, Jochem, Sander de V., Michiel, Erik W., Madeleine, Maryia, Aurélie, Mariana, Pytrik, Marjolein, Sander van D., Roelof, Sheida, Maja: it was great to have you around either for coffee, lunch or beers.

Myriam Adam was born on the 18th of October 1979 in Calais, France. She grew up with her two brothers on her parents' mixed farm in Guemps, a small rural community between Dunkerque and Calais. From 1997-2001, she attended university at the Institut Supérieur d'Agriculture in Lille, France, from which she obtained her first Master of Science degree (Ingénieur en Agriculture) in Rural Development. In 2001, she continued her studies at the College of Environmental Science and Forestry, State University of New York (CESF-SUNY), USA, where she also assisted lectures and lab sessions in human ecology and graphic engineering. As part of these studies, she worked in Ecuador in a project on the impact of land use on water quality. She obtained her second Master of Science degree in Water Management from CESF-SUNY in 2003. After returning to Europe in 2004, she was granted a Marie Curie Fellowship to work at the Agricultural Research and Education Centre Raumberg-Gumpenstein in Austria, where she carried out research for a year on nitrate leaching losses under permanent grassland, field forage and winter catch crops. In 2005, she was recruited as a research engineer at INRA-Toulouse, France, to develop a production and externalities simulator for grassland at EU scale. During this period, she became acquainted with the EU-funded project 'System for Environmental and Agricultural Modelling: Linking European Science and Society (SEAMLESS)' and the Plant Production Systems Group of Wageningen University (WU). As a follow-up of her work in Toulouse, a PhD project proposal was formulated for the development of a generic crop modelling framework for innovative cropping systems. The work presented in this thesis is the result of this project co-funded and co-supervised by the Plant Production Systems Group, Plant Research International, Wageningen University and Research Center, and the Unité Mixte de Recherche System, Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) in Montpellier, France.



1. Journal articles

- Adam M., Wery, J., Leffelaar, P.A., Ewert, F., Corbeels, M., Van Keulen, H., 2010. Simulating pea growth from wheat growth: a systematic approach for model re-assembly. *European Journal of Agronomy*, to be submitted.
- Lamanda N., Roux S., Delmotte S., Merot A., Rapidel B., Adam M., Wery J., 2010. Conceptual modelling of an agrosystem (CMA): a generic methodology and its uses in agronomy. *European Journal of Agronomy*, submitted.
- Adam, M., Belhouchette, H., Corbeels, M., Ewert, F., Perrin, A., Casellas E., Celette, F., Wery, J., 2010. Systematic evaluation to select models in a flexible and modular crop modeling framework: An application to define a model to simulate crop response to a wide range of nitrogen management. *Agricultural Systems*, submitted.
- Adam, M., Van Bussel, L.G.J., Leffelaar, P.A., Van Keulen, H., Ewert, F., 2010. Effects of modelling detail on simulated crop productivity under a wide range of climatic conditions. *Ecological Modelling*, Accepted.
- Therond, O., Hengsdijk, H., Casellas, E., Wallach, D., Adam, M., Belhouchette, H., Oomen, R., Russell, G., Ewert, F., Bergez, J.-E., Janssen, S., Wery, J., Van Ittersum, M.K., 2010. Using a cropping system model at regional scale: Low-data approaches for crop management information and model calibration. *Agriculture, Ecosystems & Environment*, In Press, Corrected Proof, Available online 17 June 2010.
- Adam, M., Ewert, F., Leffelaar, P.A., Corbeels, M., Van Keulen, H., Wery, J., 2010. CROSPAL, software that uses agronomic expert knowledge to assist modules selection for crop simulation. *Environmental Modelling & Software* 25, 946-955.
- Duru, M., Adam, M., Cruz, P., Martin, G., Ansquer, P., Ducourtieux, C., Jouany, C., Theau, J.P., Viegas, J., 2009. Modelling above-ground herbage mass for a wide range of grassland community types. *Ecological Modelling* 220, 209-225.

2. Conferences papers

- Adam, M., Leffelaar, P.A., Ewert, F., Corbeels, M., Van Keulen, H., Wery, J., 2010. Developing new crop models within a flexible crop modelling framework: use of crop physiological knowledge to change a wheat model into a pea model. Poster Presentation at XIth European Society of Agronomy congress, Montpellier, France. August 29 - September 03, 2010.
- Roux, S., Adam, M., Wery, J., 2010. On an interdisciplinary framework for the development and analysis of crop models. Oral presentation at iEMSs, International Congress on Environmental Modelling and Software, Modelling for Environment's Sake, Ottawa, Canada. July 5 – 8, 2010.
- Adam, M., Ewert, F., 2009. How to deal with modularity and flexibility in crop modeling? Oral presentation at the German Society of Agronomy. Halle, Germany. September 1-3, 2009.
- Adam, M., Belhouchette, H., Casellas, E., Therond, O., Wery, J., 2009. APES, an agricultural production and externalities simulator evaluated for two main crops in Midi-Pyrenees. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), *AgSAP conference 2009*. Egmond aan Zee, The Netherlands, 10 - 12 March, 2009. pp. 194-195.

List of Publications

- Belhouchette, H., Adam, M., Casellas, E., Celette, F., Corbeels, M., Wery, J., 2009. Performances of two crop models in various conditions: The importance of underlying assumptions. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), AgSAP conference 2009. Egmond aan Zee, The Netherlands, 10 - 12 March, 2009. pp 196-197.
- Casellas, E., Bergez, J.E., Belhouchette, H., Therond, O., Adam, M., Metay, A., Wery, J., 2009. A methodology for the evaluation and improvement of a generic biophysical Soil-Plant-Atmosphere crop model based on 'Mini-Application'. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), AgSAP conference 2009. Egmond aan Zee, The Netherlands, 10 - 12 March, 2009. pp 198-199.
- Casellas, E., Therond, O., Belhouchette, H., Wallach, D., Oomen, R.J., Russell, G., Ewert, F., Hengsdijk, H., Ittersum, M.K.v., Adam, M., Wery, J., 2009. Using a cropping system model for large scale impact assessment in Europe. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), AgSAP conference 2009. Egmond aan Zee, The Netherlands, 10 - 12 March, 2009. pp 200-201.
- Wallach, D., Evert, F.K.v., Adam, M., 2009. Parameter estimation software for crop models. In: Van Ittersum, M.K., Wolf, J., Van Laar, H.H. (Eds.), AgSAP conference 2009. Egmond aan Zee, The Netherlands, 10 - 12 March, 2009. pp 342-343.
- Adam, M., Tiftonell, P., Bergez, J.E., Ewert, F., Leffelaar, P., 2008. From detailed to summary models of the crop-soil system for larger scale applications. Oral presentation at the X ESA Congress, 15-19 September 2008, Bologna, Italy.
- Adam, M., Leffelaar, P.A., Ewert, F., Rapidel, B., Corbeels, M., Wery, J., Van Keulen, H., 2007. Developing a generic crop modelling framework: how to use expert knowledge to define crop models? Farming Systems Design 2007: an international symposium on Methodologies for Integrated Analysis of Farm Production Systems Field-farm scale design and improvement. 10-12 September 2007, The University of Catania, Sicily, Italy pp 241-242.
- Van Evert, F.K., Leffelaar, P.A., Acutis, M., Adam, M., Ewert, F., Keulen, H.v., Trevisiol, P., 2007. Practical Use of Components in Agro-Ecological Simulation, The Summer Simulation Conference Society for Modeling and Simulation, 15/18 July 2007, San Diego, CA, USA pp 55.
- Adam, M., Ewert, F., Leffelaar, P.A., Rapidel, B., Wery, J., Van Keulen, H., 2007. Development of a generic crop modelling framework. Oral presentation at the 37th Biological Systems Simulation Conference, April 17-19, 2007, Maryland, USA.

3. Reports and book chapter

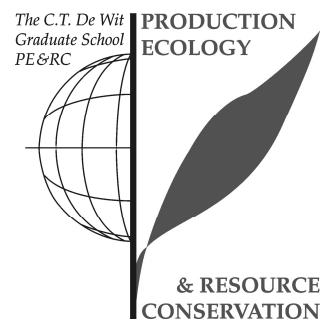
- Donatelli, M., Russell, G., Rizzoli, A.E., Acutis, M., Adam, M., et al., 2010. A component-Based Framework for Simulating Agricultural Production and Externalities, Environmental and agricultural modelling: integrated approaches for policy impact assessment. Springer, Dordrecht, pp 63-108.
- Rapidel B., Corbeels M., Adam M., 2009. Modelling options and development of the APES version for tropical-arid environments, to be implemented in the component CROPS Seamless PD 3.2. 21, 3 April 2009.

List of Publications

- Donatelli, M., Van Evert, F.K., Rutgers, B., Trevisan, M., Ewert, F., Sorce, A., Balderacchi, M., Di Guardo, A., Leffelaar, P.A., Meuter, E.C., Adam, M., Smith, J., Gottschalk, P., Acutis, M., Trevisiol, P., Gentile, A., Bellocchi, G., Carlini, L., Duru, M., Cruz, P., Gary, C., Kansou, K., Wery, J., Dupraz, C., Braudeau, E., Martin, P., 2007. Agricultural Production and Externalities Simulator (APES) prototype to be used in Prototype 1 of SEAMLESS-IF. Wageningen UR , Wageningen, pp 29.
- Donatelli, M., Rizzoli, A.E., Van Evert, F.K., Rutgers, B., Trevisan, M., Sorce, A., Balderacchi, M., Di Guardo, A., Ewert, F., Leffelaar, P.A., Meuter, E.C., Adam, M., Smith, J., Gottschalk, P., Acutis, M., Trevisiol, P., Gentile, A., Bellocchi, G., Carlini, L., Duru, M., Cruz, P., Gary, C., Kansou, K., Wery, J., Dupraz, C., Braudeau, E., Martin, P., 2007. Library of model components for process simulation relevant to production activities, Prototype 1 versions. SEAMLESS, Wageningen, pp 47.

PE&RC PhD Education Certificate

With the educational activities listed below the PhD candidate has complied with the educational requirements set by the C.T. de Wit Graduate School for Production Ecology and Resource Conservation (PE&RC) which comprises of a minimum total of 32 ECTS (= 22 weeks of activities)



Review of literature (ECTS)

- Development of a crop modelling framework

Writing of project proposal (4 ECTS)

- Development of a crop modelling framework

Post-graduate courses (8.5 ECTS)

- Art of modelling (2006)
- Scaling and governance (2009)
- Getting to the bottom of Mount Kenya; Land Dynamics (2009)

Laboratory training and working visits (3.5 ECTS)

- Crop modelling and crop physiology on pea; CIRAD, Montpellier; a number of visits throughout the PhD project (2006-2010)
- Visit of experimental sites on cotton; CIRAD, Mali, IER; 2 weeks (2007)

Competence strengthening / skills courses (2.6 ECTS)

- PhD Competence assessment (2006)
- Scientific writing (2009)
- Career orientation (2010)

PE&RC Annual meetings, seminars and the PE&RC weekend (1.5 ECTS)

- PE&RC Weekend (2007)
- PE&RC Day (2008)
- PE&RC Day (2009)

Discussion groups / local seminars / other scientific meetings (5 ECTS)

- Stats, Maths and Modelling (2007-2008)
- Spam: Spatial Method (2008-2009)

International symposia, workshops and conferences (9.8 ECTS)

- The 37th Biological Systems Simulation Conference; Maryland, USA; oral presentation (2007)
- The Farming Systems Design Conference; Catania, Sicily; poster presentation (2007)
- The X ESA Congress ; Bologna, Italy; oral presentation (2008)
- AgSAP Conference; the Netherlands; poster presentation (2009)
- The German Society of Agronomy Conference; Halle, Germany; oral presentation (2009)
- SEAMLESS/APES Meetings in Europe (2006-2009)

Lecturing / supervision of practical' s / tutorials (3 ECTS)

- Assisting Peter Leffelaar on Models for Forest and Nature Conservation; 6 days (2008)
- APES Workshops; Montpellier and Wageningen; 4 days (2009)

Super vision of 2 MSc students

- Modelling of cotton (Noemie Reboul)
- Use of APES for High price scenario study (Aurelie Perrin)

Funding

This project has been co-funded and co-supervised by the Plant Production Systems Group, Plant Research International, Wageningen University and Research Center, and the Unité Mixte de Recherche System, Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) in Montpellier, France.

This work has been carried out as part of the SEAMLESS integrated project, EU 6th Framework Programme for Research, Technological Development and Demonstration, Priority 1.1.6.3. Global Change and Ecosystems (European Commission, DG Research, Contract No. 010036-2).